

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 11:33:18 ; Search time 80.4139 Seconds
(without alignments)
3568.829 Million cell updates/sec

Title: US-10-051-909-32

Perfect score: 4075

Sequence: 1 IRSGSLAVQPTFDLDR.....PLEVITEFFAVGAKQAAKA 800

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4075	100.0	800	8	ADG47936 Corn Arab
2	3774	92.6	745	8	Adm48300 Polypepti
3	3517	86.3	747	5	Aau97201 Corn Arab
4	3517	86.3	747	6	Abu08326 Corn suga
5	3517	86.3	747	8	Adg47906 Corn Arab
6	2834	69.5	740	8	Adg47941 Arabidops
7	2674	65.6	737	5	Aau97204 Soybean A
8	2674	65.6	737	6	Abu08329 Soybean s
9	2674	65.6	737	8	Adg47912 Corn Arab
10	2584	63.4	729	5	Abb93252 Herbicida
11	2584	63.4	729	8	Adg47933 Arabidops
12	2275	55.8	734	5	Abb91078 Herbicida
13	2245	55.1	729	5	Abb92622 Herbicida
14	1670.5	41.0	486	5	Aau97205 Portion o
15	1670.5	41.0	486	6	Abu08330 Soybean A
16	1670.5	41.0	486	8	Adg47914 Soybean A
17	1424	34.9	345	5	Aau97207 Portion o
18	1424	34.9	345	6	Abu08332 Wheat sug
19	1424	34.9	345	8	Adg47918 Wheat Ara
20	943	23.1	228	5	Aau97208 Portion o
21	943	23.1	228	6	Abu08333 Wheat sug
22	943	23.1	228	8	Adg47920 Wheat Ara
23	609.5	15.0	576	7	Adc07792 Rice prot
24	562	13.8	131	5	Aau97203 Portion o
25	562	13.8	131	6	Abu08328 Rice suga

26	562	13.8	131	8	ADG47910	Adg47910 Rice Arab
27	554.5	13.6	487	6	ABG73334	Abg73334 Consensus
28	554.5	13.6	487	8	ADH51665	Adh51665 Novel hum
29	548	13.4	488	6	ABP98504	Abp98504 PFAM cons
30	548	13.4	488	6	AAE35306	Aae35306 Human sug
31	548	13.4	488	7	ADD22918	Add22918 Human sug
32	534	13.1	580	8	ADM48304	Adm48304 Polypepti
33	520.5	12.8	555	3	AAG29529	Aag29529 Arabidops
34	520.5	12.8	555	3	AAG29528	Aag29528 Arabidops
35	496	12.2	523	5	Aau97212	Aau97212 Soybean B
36	496	12.2	523	6	ABU08337	Abu08337 Soybean s
37	496	12.2	523	8	ADG47928	Adg47928 Soybean B
38	493.5	12.1	491	3	AAG32072	Aag32072 Arabidops
39	493.5	12.1	508	3	AAG32071	Aag32071 Arabidops
40	493	12.1	493	8	ADN72405	Adn72405 Thale cre
41	492	12.1	446	6	ABU43281	Abu43281 Protein e
42	489.5	12.0	513	5	Aau97210	Aau97210 Corn Beta
43	489.5	12.0	513	6	ABU08335	Abu08335 Corn suga
44	489.5	12.0	513	8	ADG47924	Adg47924 Corn Beta
45	489	12.0	529	5	Aau97214	Aau97214 Wheat sug

ALIGNMENTS

RESULT 1
ADG47936

ID ADG47936 standard; protein; 800 AA.

XX ADG47936;

DT 11-MAR-2004 (first entry)

DE Corn Arabidopsis-like sugar transport protein #2.

XX Arabidopsis-like sugar transport protein;

KW Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;

KW corn; plant.

XX Zea mays.

XX US2002199217-A1.

XX 26-DEC-2002.

PF 17-JAN-2002; 2002US-00051909.

PR 24-APR-1998; 98US-0083044P.

PR 14-APR-1999; 99US-00291922.

XX (HELE/) HELENTJARIS T G.

PI Helentjaris TG;

XX WPI; 2004-040967/04.

DR N-PSDB; ADG47935.

XX New isolated polynucleotide encoding a polypeptide having sugar transport protein activity, for producing a transformed plant and for use as probes in physical mapping.

PS Claim 18; SEQ ID NO 32; 71pp; English.

XX The invention relates to Arabidopsis-like or Beta-vulgaris-like sugar transport proteins and their corresponding nucleic acid sequences. The sequences of the invention are useful to transform a cell. These are also useful to produce a transgenic plant. Probes derived from sequences encoding sugar transport protein may be used for physical mapping. The present sequence is corn Arabidopsis-like sugar transport protein.

SQ Sequence 800 AA;

Query Match 100.0%; Score 4075; DB 8; Length 800;

Best Local Similarity 100.0%; Pred. No. 0;		Matches 800; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	IRSGSLAVQTPPTDRLRRERLLPSVIALPGLPPASCSSQEPVTSDDILEDKWGA	60
Db	1	IRSGSLAVQTPPTDRLRRERLLPSVIALPGLPPASCSSQEPVTSDDILEDKWGA	60
QY	61	LVAIVASIGNLQGDNDATAAVALYIKKEFQONEPTVEGLIVSMISLIGATVITFSGP	120
Db	61	LVAIVASIGNLQGDNDATAAVALYIKKEFQONEPTVEGLIVSMISLIGATVITFSGP	120
QY	121	LSDSICRRPMLILSSILYFSGGLIMLSPNVYVLLARFVDGFGICLAVTLVPLVISIA	180
Db	121	LSDSICRRPMLILSSILYFSGGLIMLSPNVYVLLARFVDGFGICLAVTLVPLVISIA	180
QY	181	PSIRGLLNTLPQFSGGGMFLSYCMVFGMSLSPSPDNRIMLGVLAIPLSPFFGLTIFYL	240
Db	181	PSIRGLLNTLPQFSGGGMFLSYCMVFGMSLSPSPDNRIMLGVLAIPLSPFFGLTIFYL	240
QY	241	PSPRWLVSKGRMAEAKKVLQKRGKDDVSGLSLLEGLVGGDTSEIYIIGPATAA	300
Db	241	PSPRWLVSKGRMAEAKKVLQKRGKDDVSGLSLLEGLVGGDTSEIYIIGPATAA	300
QY	301	DLVATDGDREQITLYGPEBQSWIARPSKPIMLGSLVSLASRHGSMWNQSVPLMDPIVT	360
Db	301	DLVATDGDREQITLYGPEBQSWIARPSKPIMLGSLVSLASRHGSMWNQSVPLMDPIVT	360
QY	361	LFSGVHNNPQAGGSMRSTLFPNFGSMFVTQDHAQNEQWDEENLHRDEEYASDAGGD	420
Db	361	LFSGVHNNPQAGGSMRSTLFPNFGSMFVTQDHAQNEQWDEENLHRDEEYASDAGGD	420
QY	421	YEDNLHSPILSRQATCAEKGDIIVHGHRSALSMRQTLIGBGGGVSTDIIGGQWLAW	480
Db	421	YEDNLHSPILSRQATCAEKGDIIVHGHRSALSMRQTLIGBGGGVSTDIIGGQWLAW	480
QY	481	KWSEKEGNGRKEGFKRYLHQEGVPGSRGSIIVSLPGGGDVFGSGFFVHAALVSQSA	540
Db	481	KWSEKEGNGRKEGFKRYLHQEGVPGSRGSIIVSLPGGGDVFGSGFFVHAALVSQSA	540
QY	541	LFSGKGLAEPRMDAAMVHSEVAAGSRWKDLFEPGVRALLVGVGIQLOQFAGINGVL	600
Db	541	LFSGKGLAEPRMDAAMVHSEVAAGSRWKDLFEPGVRALLVGVGIQLOQFAGINGVL	600
QY	601	YTPQILEQAGVAVILSKFGLSSASASILISLTLTLLMPCIGFANLMDLSGRFFLLG	660
Db	601	YTPQILEQAGVAVILSKFGLSSASASILISLTLTLLMPCIGFANLMDLSGRFFLLG	660
QY	661	TIPILIASIVLVNLDLGLTALHALLSTSVIYVFCFVWGFGPIPNILCBIFFPRV	720
Db	661	TIPILIASIVLVNLDLGLTALHALLSTSVIYVFCFVWGFGPIPNILCBIFFPRV	720
QY	721	RGLCIAICAFITWIGDIIIVTYSPLVMAIGAIGVFSIYAVVCLISFVFLKVPETKGM	780
Db	721	RGLCIAICAFITWIGDIIIVTYSPLVMAIGAIGVFSIYAVVCLISFVFLKVPETKGM	780
QY	781	PLEVITEPPAVGAQAANA 800	
Db	781	PLEVITEPPAVGAQAANA 800	

RESULT 2
ADM48300
ID ADM48300 standard; protein; 745 AA.
AC
XX
XX
XX
DT 03-JUN-2004 (first entry)
DE Polypeptide sequence #350 useful in producing transgenic plants.
XX
XX Plant; maize; rice; sorghum; thale cress; soybean; wheat; transgenic;
KW osmotic stress; sugar transport; cell cycle pathway; plant height;
KW carbohydrate transport; crop productivity; plant growth;

stress resistance; disease resistance; insect resistance; heat tolerance; nitrogen assimilation; water stress tolerance; photosynthetic carbon fixation; virus resistance; gene therapy.
Zea mays.
US2003233670-A1.
18-DEC-2003.
04-DEC-2002; 2002US-00310154.
04-DEC-2001; 2001US-0337358P.
(EDGE/) EDGERTON M D.
(CHOM/) CHOMET P S.
(LACC/) LACCETTI L B.
Edgerton MD, Chomet PS, Laccetti LB;
WPI; 2004-061374/06.
N-PSDB; ADM47932.
New polynucleotide, useful for manipulating plant protein quality, improving plant growth, yield and crop productivity or grain composition or producing plants with improved properties.
Claim 8; SEQ ID NO 718; 144pp; English.
The present invention relates to polynucleotide sequences, and the proteins they encode. The sequences are isolated from a variety of organisms such as plants (e.g. maize, rice, sorghum, thale cress, soybean, and wheat), cyanobacteria, bacteria, yeast and other fungi. The polynucleotide and polypeptide sequences of the invention are useful in the production of transgenic plants that have improved properties. Also disclosed are methods of producing fertile transgenic plants, preferably maize, with desired phenotypes. The polynucleotide and polypeptide sequences are useful for improving plants by providing protection against osmotic stress, improving altering sugar transport and/or metabolism, modifying the cell cycle pathway, reducing plant height, modifying carbohydrate transport, improving crop productivity, improving plant growth and stress resistance, improving disease resistance, improving insect resistance, improving cold or heat tolerance, improving nitrogen assimilation, improving stalk strength, improving water stress tolerance, improving photosynthetic carbon fixation, improving biotic and abiotic stress resistance, improving resistance to oxidative stress, providing increased vigour, reducing senescence, and conferring virus resistance. The present sequence represents a polypeptide sequence of the invention. Note: The sequence data for this patent is not provided in the printed specification but is obtained in electronic format from the USPTO website at segdata.uspto.gov.

Query Match 32.6%; Score 3774; DB 8; Length 745;		Best Local Similarity 99.8%; Pred. No. 0;		Matches 742; Conservative 1; Mismatches 2; Indels 0; Gaps 0;	
QY	56	MSGAVLVAIVASIGNLQGDNDATAAVALYIKKEFQONEPTVEGLIVSMISLIGATVIT	115		
Db	1	MSGAVLVAIVASIGNLQGDNDATAAVALYIKKEFQONEPTVEGLIVSMISLIGATVIT	60		
QY	116	TFSGPLSDSIRGRPMILSSILYFSGGLIMLSPNVYVLLARFVDGFGICLAVTLVPLV	175		
Db	61	TFSGPLSDSIRGRPMILSSILYFSGGLIMLSPNVYVLLARFVDGFGICLAVTLVPLV	120		
QY	176	ISBIAPEIRGLLNTLPQFSGGGMFLSYCMVFGMSLSPSPDNRIMLGVLAIPLSPFFGL	235		
Db	121	ISBIAPEIRGLLNTLPQFSGGGMFLSYCMVFGMSLSPSPDNRIMLGVLAIPLSPFFGL	180		
QY	236	TIFYPESPRWLVSKGRMAEAKKVLQKRGKDDVSGLSLLEGLVGGDTSEIYIIGP	295		
Db	181	TIFYPESPRWLVSKGRMAEAKKVLQKRGKDDVSGLSLLEGLVGGDTSEIYIIGP	240		

QY	296	ATEAADDLVTDGKEQITLYGPEEQSWIARPSKGPIMLGSVLSLASHGSMVNSVPLM	355
Db	241	ATEAADDLVTDGKEQITLYGPEEQSWIARPSKGPIMLGSVLSLASHGSMVNSVPLM	300
QY	356	DPIVTLFGSVHENPQAGGSMRSTLFPNFGSMFVTDQHAKEQNDENLHRDDEEYASD	415
Db	301	DPIVTLFGSVHENPQAGGSMRSTLFPNFGSMFVTDQHAKEQNDENLHRDDEEYASD	360
QY	416	GAGGDIYEDNLHSPILLSROATGAEGKDI VHHGHRGSALSMMRQTLLEGCGDVSTDIGG	475
Db	361	GAGGDIYEDNLHSPILLSROATGAEGKDI VHHGHRGSALSMMRQTLLEGCGDVSTDIGG	420
QY	476	WQLAWKSEKEGNGRKEGGFRVYLHOEGVPGSRGSIIVSLPGGDPVFESEFVHAAL	535
Db	421	WQLAWKSEKEGNGRKEGGFRVYLHOEGVPGSRGSIIVSLPGGDPVFESEFVHAAL	480
QY	536	VSQSALFSGKLAEPMSDAAMVHSEVAAGSRWKDLFEPGVRRALLVGVGIQLQOFAG	595
Db	481	VSQSALFSGKLAEPMSDAAMVHSEVAAGSRWKDLFEPGVRRALLVGVGIQLQOFAG	540
QY	596	INGVLYTPQILEQAGAVIILSKFGLSSASASILLISLTLMLPCTGFAMLLMDLSGRR	655
Db	541	INGVLYTPQILEQAGAVIILSKFGLSSASASILLISLTLMLPCTGFAMLLMDLSGRR	600
QY	656	FLLLGTIPILIASILVSVNLIDGLTIAHALLSTVSVVVFCCFVVGFGPIPNILCAEI	715
Db	601	FLLLGTIPILIASILVSVNLIDGLTIAHALLSTVSVVVFCCFVVGFGPIPNILCAEI	660
QY	716	FPTRVRGLCIAICAFTFWIGDIIVTYSLPVNLNAGLAGVFSIYAVVCLISFVFVLKVP	775
Db	661	FPTRVRGLCIAICAFTFWIGDIIVTYSLPVNLNAGLAGVFSIYAVVCLISFVFVLKVP	720
QY	776	ETKGNPLEVITEFFAVGAKQAAKA	800
Db	721	ETKGNPLEVITEFFAVGAKQAAKA	745
RESULT 3			
AAU97201			
ID	AAU97201	standard; protein; 747 AA.	
AC	AAU97201;		
XX			
DT	27-AUG-2002	(first entry)	
XX			
DE	Corn Arabidopsis thaliana-like sugar transport protein.		
XX			
XW	Corn; Arabidopsis thaliana-like sugar transport protein;		
XW	carbohydrate transport; grain filling; annual field crop; plant.		
XX			
OS	Zea mays.		
XX			
FH	Key	Location/Qualifiers	
FT	Misc-difference 129	/label= Unknown	
FT	Misc-difference 133..134	/label= Unknown	
FT	Misc-difference 144	/label= Unknown	
FT	Misc-difference 178	/label= Unknown	
FT	Misc-difference 207	/label= Unknown	
FT	Misc-difference 218	/label= Unknown	
FT	Misc-difference 220	/label= Unknown	
FT	Misc-difference 236	/label= Unknown	
FT	Misc-difference 236	/label= Unknown	
PN	US6383776-B1.		
XX			

PD	07-MAY-2002.		
XX	14-APR-1999;	99US-00291922.	
XX	24-APR-1998;	98US-0083044P.	
XX	(DUPO) DU PONT DE NEMOURS & CO E I.		
PI	Allen SM, Hitz WD, Kinney AJ, Tingey SV;		
DR	WPI; 2002-453364/48.		
DR	N-PSDB; ABK51962.		
XX	New nucleic acid encoding plant sugar-transport proteins, useful for		
PT	preparing transgenic plants with altered carbohydrate distribution.		
XX	Claim 1; Fig 1; 54pp; English.		
XX	The present invention relates to the isolation of plant polynucleotide		
CC	sequences encoding an Arabidopsis thaliana-like sugar transport protein		
CC	or Beta vulgaris-like sugar transport protein. The polynucleotide		
CC	sequences are useful for altering the level of sugar transport proteins		
CC	in plants, i.e. for control of carbohydrate transport and distribution in		
CC	plant cells, e.g. during grain filling of annual field crops (e.g. corn,		
CC	rice, soybeans, and wheat), and, for studying carbohydrate flows and		
CC	sugar transport. The polynucleotide sequences can also be used to isolate		
CC	cDNA sequences and genes that encode homologues of the new proteins. The		
CC	present sequence represents a corn Arabidopsis thaliana-like sugar		
CC	transport protein		
XX	Sequence 747 AA;		
SQ	Query Match	86.3%; Score 3517; DB 5; Length 747;	
	Best Local Similarity	91.6%; Pred. No. 0;	
	Matches 684; Conservative 27; Mismatches 34; Indels 2; Gaps 2;		
QY	56	MSGAVLVAIVASIGNLLQGDNDATIAAVALYIKKEFQIQNEPTVEGLIYVMSLIGATIVT	115
Db	1	MCGAVMVAIAASIGNLLQGDNDATIAAVALYIKKEFNLOSEPIEGLIVAMFUGATVIT	60
QY	116	TFSGPLSDSISGRPRMLILSILYFFSGLIWMLSPNYVLLARFVDFGFIGLAVTLVPLY	175
Db	61	TSPGPRADCVGRRPMLVASAVLYFVGLVWLWAPIVYILLARLIDGFIGLAVTLVPLY	120
QY	176	ISEIAPSEIR-GLLNTLPQFSG-SGGMFLSYCMVFGMSLSPSPDWIRMLGLVAIPSLFFP	233
Db	121	ISETAPHRXGWGXNTLPQFIVGVGMFLSYCMVFGMSLMPKPDWIRMLGLVSLPSLYF	180
QY	234	GLTIFVLPESPRWLKSGMAEAKVYLQIRGKDDVSGELSLLEGVGGDSIEEYII	293
Db	181	GLTVFVLPESPRWLKSGMAEAKRVXQRLRGREDVSXEXALLVEGLVGKDTRIEYII	240
QY	294	GPATEAADLVTDGKEQITLYGPEEQSWIARPSKGPIMLGSVLSLASHGSMVNSVSP	353
Db	241	GPATEAADLVTDGKEQITLYGPEEQSWIARPSKGPIMLGSVLSLASHGSMVNSVSP	300
QY	354	LMDPIVTLFGSVHENPQAGGSMRSTLFPNFGSMFVTDQHAKEQNDENLHRDDEEYA	413
Db	301	LMDPIVTLFGSVHENPQAGGSMRSTLFPNFGSMFVTDQHAKEQNDENLHRDDEEYA	360
QY	414	SDGAGDIYEDNLHSPILLSROATGAEGKDI VHHGHRGSALSMMRQTLLEGCGDVSTDIG	473
Db	361	SDGAGDIYEDNLHSPILLSROATGAEGKDI VHHGHRGSALSMMRQTLLEGCGDVSTDIG	420
QY	474	GGWQLAWKSEKEGNGRKEGGFRVYLHOEGVPGSRGSIIVSLPGGDPVFESEFVHAAL	533
Db	421	GGWQLAWKSEKEGNGRKEGGFRVYLHOEGVPGSRGSIIVSLPGGDPVFESEFVHAAL	480
QY	534	ALYSQSALFSGKLAEPMSDAAMVHSEVAAGSRWKDLFEPGVRRALLVGVGIQLQOF	593
Db	481	ALYSQSALFSGKLAEPMSDAAMVHSEVAAGSRWKDLFEPGVRRALLVGVGIQLQOF	540
QY	594	AGINGVLYTTPQILEQAGAVIILSKFGLSSASASILLISLTLMLPCTGFAMLLMDLSG	653

Db 541 AGINGVLYTQILQAGVAVILSKFGLSSASILSSLTLLMLPCIGFAMLLMDLSG 600
 QY 654 RRFLLGTIPILIASLVILVSNLIDLTGLAHALLSTVSVIVFCFCFVWGSPINILCA 713
 Db 601 RRFLLGTIPILIASLVILVSNLIDLTGLAHALLSTVSVIVFCFCFVWGSPINILCA 660
 QY 714 EIFPTRVRGLCIAICAFTHWIGDIIIVTSLPMLNAIGLAGVFSIYAVVCLISFVFVLK 773
 Db 661 EIFPTRVRGLCIAICAFTHWIGDIIIVTSLPMLNAIGLAGVFSIYAVVCLISFVFVLK 720
 QY 774 VPETKGMPLVITEFFAVGAKQAARAKA 800
 Db 721 VPETKGMPLVITEFFAVGAKQAARAKA 747

RESULT 4
 ABU08326
 ID ABU08326 standard; protein; 747 AA.
 AC ABU08326;
 DT 29-MAY-2003 (first entry)
 DE Corn sugar transport protein #1.
 KW Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
 KW plant sugar transport protein; carbohydrate transport; soybean;
 KW carbohydrate distribution; plant.
 OS Zea mays.
 PN US2002178468-A1.
 PD 28-NOV-2002.
 PF 17-JAN-2002; 2002US-00051902.
 PR 24-APR-1998; 98US-0083044P.
 PR 14-APR-1999; 99US-00291922.
 PA (ALDE/) ALLEN S. M.
 PA (HITZ/) HITZ W. D.
 PA (KINN/) KINNEY A. J.
 PA (TING/) TINGEY S. V.
 PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
 DR WPI; 2003-340957/32.
 DR N-PSDB; ABX93198.
 XX Novel plant sugar transport proteins and nucleic acid encoding the
 PT protein useful for producing transgenic plants having altered levels of
 PT sugar transport protein.
 PS Claim 5; Fig 1; 56pp; English.
 CC The present invention relates to the isolation of Arabidopsis thaliana-
 CC like or Beta vulgaris-like sugar transport proteins, and the
 CC polynucleotide sequences encoding them. The plant sugar transport
 CC proteins of the invention have been isolated from corn, rice, soybean,
 CC and wheat. The polypeptides of the invention may be used for altering the
 CC level of expression of a sugar transport protein in a host cell, by
 CC transforming a host cell with a chimeric construct encoding all, or a
 CC portion of the sugar transport protein, in sense or antisense
 CC orientation. Particularly, the polypeptides may provide a means to
 CC control carbohydrate transport and distribution in plants. ABU08326-
 CC ABU08333 represent Arabidopsis thaliana-like sugar transport proteins
 XX Sequence 747 AA;
 SQ

Query Match 86.3%; Score 3517; DB 6; Length 747;
 Best Local Similarity 91.6%; Pred. No. 0;

Matches 684; Conservative 27; Mismatches 34; Indels 2; Gaps 2;
 QY 56 MSGAVLVAVASIGNLQGNWNTAAAVLYIKKFFOLQNEBTVGLIVSNLSLICATVIT 115
 Db 1 MGAVMVAIAASIGNLQGNWNTAGAVLYIKKFNLOSEPLIEGLIVAMFLICATVIT 60
 QY 116 TFSGPLSDSIGRRPMLILSSIIYFFSGTLMSPNVYLLARFVDFGFGIGLAVTLVPLY 175
 Db 61 TSPGPRADCVGRPMVASAVLYFVSLVMLWAPIVYILLARLIDGFGIGLAVTLVPLY 120
 QY 176 ISEIAPSEIR-GLLNTLPQPSG-SCGMFLSYCNWFGMSLSPSPDRIMGLVLAIPSLFF 233
 Db 121 ISETAPHRXSWGXNTLPQFIGVXGMFLSYCNWFGMSLSPSPDRIMGLVLAIPSLFF 180
 QY 234 GLTIFVLPESPRWLVSKGRMAEAKVLOKLRGKDDVSGELSLLEGLFVGGDTSEEEYII 293
 Db 181 GLTIFVLPESPRWLVSKGRMAEAKVXQRLRGREDVSEXALLVGLGVGKDTRIEXEYII 240
 QY 294 GPATEAADLVTDGKEQITLYGPEEGOSWIARPSKGPIMGLSVLSASRHSWVNSQVSP 353
 Db 241 GPATEAADLVTDGKEQITLYGPEEGOSWIARPSKGPIMGLSVLSASRHSWVNSQVSP 300
 QY 354 LNDPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFVTDQAHKNEQWDEENLHRDDEYA 413
 Db 301 LNDPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFVTDQAHKNEQWDEENLHRDDEYA 360
 QY 414 SDGAGGYEDNLHSPLLSRQATGAEKDIIVHGHRSALSRRQTLLEGSGDVSSTDIG 473
 Db 361 SDGAGGYEDNLHSPLLSRQATGAEKDIIVHGHRSALSRRQTLLEGSGDVSSTDIG 420
 QY 474 GGMOLAKWSEKEGNGRKEGFKRVIYHOGVPGSRGSIIVSLPGGDVPEGSEFVHAA 533
 Db 421 GGMOLAKWSEKEGNGRKEGFKRVIYHOGVPGSRGSIIVSLPGGDVPEGSEFVHAA 480
 QY 534 ALVQSALFSGLABPRMSDAAMVHPSEVAAKSGSKDLFFPGVRRALLVGVGIQLQOF 593
 Db 481 ALVQSALFSGLABPRMSDAAMVHPSEVAAKSGSKDLFFPGVRRALLVGVGIQLQOF 540
 QY 594 AGINGVLYTQILQAGVAVILSKFGLSSASILSSLTLLMLPCIGFAMLLMDLSG 653
 Db 541 AGINGVLYTQILQAGVAVILSKFGLSSASILSSLTLLMLPCIGFAMLLMDLSG 600
 QY 654 RRFLLGTIPILIASLVILVSNLIDLTGLAHALLSTVSVIVFCFCFVWGSPINILCA 713
 Db 601 RRFLLGTIPILIASLVILVSNLIDLTGLAHALLSTVSVIVFCFCFVWGSPINILCA 660
 QY 714 EIFPTRVRGLCIAICAFTHWIGDIIIVTSLPMLNAIGLAGVFSIYAVVCLISFVFVLK 773
 Db 661 EIFPTRVRGLCIAICAFTHWIGDIIIVTSLPMLNAIGLAGVFSIYAVVCLISFVFVLK 720
 QY 774 VPETKGMPLVITEFFAVGAKQAARAKA 800
 Db 721 VPETKGMPLVITEFFAVGAKQAARAKA 747
 RESULT 5
 ADG47906
 ID ADG47906 standard; protein; 747 AA.
 AC ADG47906;
 XX 11-MAR-2004 (first entry)
 DT Corn Arabidopsis-like sugar transport protein #1.
 DE Arabidopsis-like sugar transport protein;
 KW Arabidopsis-like sugar transport protein;
 KW Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;
 KW corn; plant.
 OS Zea mays.
 PF Key Location/Qualifiers
 FT Misc-difference 129

FT		/note= "Encoded by NAT"	
FT	Misc-difference 133	/note= "Encoded by TGN"	
FT	Misc-difference 134	/note= "Encoded by TNG"	
FT	Misc-difference 144	/note= "Encoded by AGN"	
FT	Misc-difference 178	/note= "Encoded by ATN"	
FT	Misc-difference 195	/note= "Encoded by GTN"	
FT	Misc-difference 207	/note= "Encoded by NTG"	
FT	Misc-difference 218	/note= "Encoded by NGG"	
FT	Misc-difference 220	/note= "Encoded by ANG"	
FT	Misc-difference 236	/note= "Encoded by TNA"	
XX	US2002199217-A1.		
XX	26-DEC-2002.		
XX	17-JAN-2002; 2002US-00051909.		
XX	24-APR-1998; 98US-0083044P.		
XX	14-APR-1999; 99US-00291922.		
XX	(HELE/) HELENTJARIS T G.		
XX	Helentjaris TG;		
XX	WPI; 2004-040967/04.		
XX	N-PSDB; ADG47905.		
XX	New isolated polynucleotide encoding a polypeptide having sugar transport protein activity; for producing a transformed plant and for use as probes in physical mapping.		
XX	Example 3; SEQ ID NO 2; 71pp; English.		
XX	The invention relates to Arabidopsis-like or Beta-vulgaris-like sugar transport proteins and their corresponding nucleic acid sequences. The sequences of the invention are useful to transform a cell. These are also useful to produce a transgenic plant. Probes derived from sequences encoding sugar transport protein may be used for physical mapping. The present sequence is corn Arabidopsis-like sugar transport protein.		
XX	Sequence 747 AA;		
SQ	Query Match 86.3%; Score 3517; DB 8; Length 747;		
	Best Local Similarity 91.6%; Pred. No. 0;		
	Matches 684; Conservative 27; Mismatches 34; Indels 2; Gaps 2;		
QY	56 MSGAVLVAIVASIGNLLQGDNDATIAAVALYIKKEFQIQNEPTVEGLIVSMISLIGATIVT 115		
DB	1 MCGAVVAIVASIGNLLQGDNDATIAAGAVLYIKKEFNQSEPLIEGLIVAMFLIGATVIT 60		
QY	116 TSPGFLSDSIGRRPMLISSILYFFSGILMLWSNNVYLLARFVDGFGIGLAVTLVPLY 175		
DB	61 TSPGPRADCVGRPRMLVASAVLYFVSGVLMWAPIVYILLARLIDGFGIGLAVTLVPLY 120		
QY	176 ISEIAPSEIR-GLLNTLPQFSG-SCGMFLSYCMVFCMSLSPDWRIMLGVLAIPLSPFF 233		
DB	121 ISETAPHRXGWXXNTLPQFICVXGFMFLSYCMVFCMSLMPKPDWRLMLGVLSPLSYF 180		
QY	234 GLTIFYLPSPRWLVSKGRMAFAKVLQKLRGKDDVSGELSLLEGLVEVGGDTSEIEVII 293		
DB	181 GLTIFYLPSPRWLVSKGRMAFAKVLQKLRGREDVSVEXALLVEGLVGKDXEYII 240		
QY	294 GPATEAADLVTDGKEQITLYGPEGQSWIARPSKGPIMLGSVLSLASRHGSMWNQSPV 353		
DB	241 GPATEAADLVTDGKEQITLYGPEGQSWIARPSKGPIMLGSVLSLASRHGSMWNQSPV 300		
QY	354 LMDPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFVTDQHAQNEQWDEENLHRDDEYA 413		
DB	301 LMDPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFVTDQHAQNEQWDEENLHRDDEYA 360		
QY	414 SDGAGGDYEDNLHSPLLSQATGAEKDIYVHHGHSALSMSEROTLLGEGGDVGSSTDIG 473		
DB	361 SDGAGGDYEDNLHSPLLSQATGAEKDIYVHHGHSALSMSERQSLGEGGDVGSSTDIG 420		
QY	474 GGQWLAKWSEKEGNGRKEGGFKRVYLHOGVPGSRGSIIVSLPGGGDVFFGSEFVHAA 533		
DB	421 GGQWLAKWSEKEGNGRKEGGFKRVYLHOGVPGSRGSIIVSLPGGGDVFFGSEFVHAA 480		
QY	534 ALVQSALFSKGLAEPRMSDAAMVHPSEVAAGKSWKDLFEPGVRALLVGVGIQLQOF 593		
DB	481 ALVQSALFSKGLAEPRMSDAAMVHPSEVAAGKSWKDLFEPGVRALLVGVGIQLQOF 540		
QY	594 AGINGVLYYTPQILEOAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLLMDLSG 653		
DB	541 AGINGVLYYTPQILEOAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLLMDLSG 600		
QY	654 RRFLLLGTIPILIASLVILVWSNLIDLGLTAHALISTVSVIVYFCFVNGFGPIPNILCA 713		
DB	601 RRFLLLGTIPILIASLVILVWSNLIDLGLTAHALISTVSVIVYFCFVNGFGPIPNILCA 660		
QY	714 EIPFTRVRLGCTAICAFTEWIGDIIIVTYSIPYMLNAIGLAGVFSIYAVVCLISFVFVLK 773		
DB	661 EIPFTRVRLGCTAICAFTEWIGDIIIVTYSIPYMLNAIGLAGVFSIYAVVCLISFVFVLK 720		
QY	774 VPETKGMPLFVITEFFAVGAKQAQAAKA 800		
DB	721 VPETKGMPLFVITEFFAVGAKQAQAAKA 747		
XX	RESULT 6		
XX	ADG47941		
XX	ID ADG47941 standard; protein; 740 AA.		
XX	AC ADG47941;		
XX	XX 11-MAR-2004 (first entry)		
XX	Arabidopsis thaliana-like sugar transport protein #2.		
XX	Arabidopsis-like sugar transport protein;		
XX	Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;		
XX	mouse-ear cross; plant.		
XX	Arabidopsis thaliana.		
XX	US2002199217-A1.		
XX	26-DEC-2002.		
XX	17-JAN-2002; 2002US-00051909.		
XX	24-APR-1998; 98US-0083044P.		
XX	14-APR-1999; 99US-00291922.		
XX	(HELE/) HELENTJARIS T G.		
XX	Helentjaris TG;		
XX	WPI; 2004-040967/04.		
XX	New isolated polynucleotide encoding a polypeptide having sugar transport protein activity; for producing a transformed plant and for use as probes in physical mapping.		
XX	Example 3; SEQ ID NO 37; 71pp; English.		
XX	The invention relates to Arabidopsis-like or Beta-vulgaris-like sugar		

CC transport proteins and their corresponding nucleic acid sequences. The
 CC sequences of the invention are useful to transform a cell. These are also
 CC useful to produce a transgenic plant. Probes derived from sequences
 CC encoding sugar transport protein may be used for physical mapping. The
 CC present sequence is wheat Arabidopsis thaliana-like sugar transport
 CC protein.
 XX
 SQ Sequence 740 AA;

Query Match 69.5%; Score 2834; DB 8; Length 740;
 Best Local Similarity 74.5%; Pred. No. 1e-265;
 Matches 553; Conservative 75; Mismatches 108; Indels 6; Gaps 5;
 QY 56 MSGAVLVAIVASIGNLLQGNWNTAAAVLYIKKEFQNLQNEPTVEGLVMSLIGATVIT 115
 DB 1 MAGAVLVAIVASIGNLLQGNWNTAAAGAVLYIKKEFNQSEPLIEGLIVAMSLIGATIT 60
 QY 116 TFSGLSDSISGRPRMLILSSILYFSGGLTLMWSPNVYLLARFVDFGIGLAVTLVPLV 175
 DB 61 TFSGAVDSFGRRPMLIASAVLYFVSGGLVMAWNVYLLARLIDGFGIGLAVTLVPLV 120
 QY 176 ISEIAPSEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPSPDWMRLMGLVLAIPSLFFGL 235
 DB 121 ISETAPTDIRGLNLTLPQFSGSGMFLSYCMVFGMSLMPQDWMRLMGLVLSIPSLIFAL 180
 QY 236 TIFYPESPRLVSKGRMAEKVLOKRGDVSSELSLLEGLVGGDTISIEYIIGP 295
 DB 181 TIFYPESPRLVSKGRMAEKVLOKRGDVSSELSLLEGLVGGDTISIEYIIGP 240
 QY 296 ATEAADDLVTDGKEQITLYGPEEQSWIARPSKIPMLGSLVLSLRHSGMWNQSVPLM 355
 DB 241 DDELADEGLAP-DPEKIKLYGPEELSWARPVHGQSLGALGLISRHSGMWSQGPLV 299
 QY 356 DPVTLFSGVHNNPQAGGVRSTLPFPNGSMFVSTQDHAQXNEQWDEENLHRDDEYASD 415
 DB 300 DPVTLFSGVHNNPQAGGVRSTLPFPNGSMFVSTQDHAQXNEQWDEENLHRDDEYASD 358
 QY 416 GAGGVDYENLHSPILSRQATGAEGKDI-VHGHGRGSLSMEROTLLGEGGVSSTDIIG 474
 DB 359 HGGDDIEDSLQSLISROATSVGEKETAAPHGSMGAVG--RSSSLMQGGEAVSSMGIGG 416
 QY 475 GWQLAWKWSKEGNGRKEGQFKRYVYHQGVPGSRGSIIVSLPGGDFEGSEFVHAAA 534
 DB 417 GWQLAWKWTREGADGEKEGGFQRIYLHEEGVTGDRRGSILSLP-GGDPVPGGFEVQAAA 475
 QY 535 LVQSALFSGKLAPRMSDRAWHPSEVAAGSRKWLPEFGVRRALLVGVGQILQOFA 594
 DB 476 LVQSALYSKELMEORLAGPAPWHPSOVAKGPKWADLFEFGVHALFVGIGQILQOFA 535
 QY 595 GINGVLYYTPQILEQAGVAVILSKFGLSSASASILISLTLTLLMPCIGFAMLIMDSGR 654
 DB 536 GINGVLYYTPQILEQAGVGVLLANIGLSSASILISLTLTLLMPSIGIAMRLMDMSGR 595
 QY 655 RFLLTGTPILIASVILVNSLIDGLTALHALLSTVSVYFCFVWGPGPIPNILCAE 714
 DB 596 RFLLTATPILIVAILILNLDVGTVMWASLSTSVILYFCFVWGPGPIPNILCAE 655
 QY 715 IFFTFRVGLCAICAFITFWIGDIITVSLPWLNAIGLAGVFSIYAVVCLISFVFLKV 774
 DB 656 IFFTFRVGLCAICAFITFWIGDIITVSLPWLNAIGLAGVFSIYAVVCLISFVFLKV 715
 QY 775 PETKGMPLVETTFEFAVGAQOA 796
 DB 716 PETKGMPLVETTFEFAVGAQOA 737

RESULT 7
 AAU97204
 ID AAU97204 standard; protein; 737 AA.
 XX
 AC AAU97204;
 XX
 DT 27-AUG-2002 (first entry)

XX Soybean Arabidopsis thaliana-like sugar transport protein.
 DE
 XX Soybean; Arabidopsis thaliana-like sugar transport protein;
 KW carbohydrate transport; grain filling; annual field crop; plant.
 KW
 XX Glycine max.
 OS
 XX US6383776-B1.
 PN
 XX 07-MAY-2002.
 PD
 XX 14-APR-1999; 99US-00291922.
 PF
 XX 24-APR-1998; 98US-0083044P.
 PR
 XX (DUPO) DU PONT DE NEWMOURS & CO E I.
 PA
 XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;
 PI
 XX WPI; 2002-453364/48.
 XX N-PSDB; ABX51965.
 DR
 XX New nucleic acid encoding plant sugar-transport proteins, useful for
 PT preparing transgenic plants with altered carbohydrate distribution.
 PS
 XX Claim 1; Fig 1; 5app; English.
 XX The present invention relates to the isolation of plant polynucleotide
 CC sequences encoding an Arabidopsis thaliana-like sugar transport protein
 CC or Beta vulgaris-like sugar transport protein. The polynucleotide
 CC sequences are useful for altering the level of sugar transport proteins
 CC in plants, i.e. for control of carbohydrate transport and distribution in
 CC plant cells, e.g. during grain filling of annual field crops (e.g. corn,
 CC rice, soybeans, and wheat), and, for studying carbohydrate flows and
 CC sugar transport. The polynucleotide sequences can also be used to isolate
 CC cDNA sequences and genes that encode homologues of the new proteins. The
 CC present sequence represents a soybean Arabidopsis thaliana-like sugar
 CC transport protein
 XX
 SQ Sequence 737 AA;

Query Match 65.6%; Score 2674; DB 5; Length 737;
 Best Local Similarity 69.9%; Pred. No. 3.8e-250;
 Matches 522; Conservative 92; Mismatches 115; Indels 18; Gaps 8;
 QY 56 MSGAVLVAIVASIGNLLQGNWNTAAAVLYIKKEFQNLQNEPTVEGLVMSLIGATVIT 115
 DB 1 MKGAVLVAIVASIGNFLQGNWNTAAGANGYIKKDLALGT--TWERLVVGSMLIGATVIT 58
 QY 116 TFSGLSDSISGRPRMLILSSILYFSGGLTLMWSPNVYLLARFVDFGIGLAVTLVPLV 175
 DB 59 TCSGPIADWLGRPRMIISSVLYFLGLVLMWSPNVYLLARLDDGFGIGLAVTLVPLV 118
 QY 176 ISEIAPSEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPSPDWMRLMGLVLAIPSLFFGL 235
 DB 119 ISETAPSEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPSPDWMRLMGLVLAIPSLFFGL 178
 QY 236 TIFYPESPRLVSKGRMAEKVLOKRGDVSSELSLLEGLVGGDTISIEYIIGP 295
 DB 179 TIFYPESPRLVSKGRMAEKVLOKRGDVSSELSLLEGLVGGDTISIEYIIGP 238
 QY 296 ATEAADDLVTDGKEQITLYGPEEQSWIARPSKIPMLGSLVLSLRHSGMWNQSVPLM 355
 DB 239 ADDVADGHEHATEKDKIRLYGSQAGLSWLSKFTVQSSIG---LASHGSIINQSMPLM 294
 QY 356 DPVTLFSGVHNNPQAG--GSMRSTLPFPNGSMFVSTQDHAQXNEQWDEENLHRDDEYASD 413
 DB 295 DPVTLFSGVHNNPQAG--GSMRSTLPFPNGSMFVSTQDHAQXNEQWDEENLHRDDEYASD 354
 QY 414 SDGAGGVDYENLHSPILSRQATGAEGKDI-VHGHGRGSLSMEROTLLGEGGVSSTDIIG 471
 DB 355 SDATRGSDNDNLHSPILSRQATGAEGKDI-VHGHGRGSLSMEROTLLGEGGVSSTDIIG 413

472 IGGWQLAWKWSKEGNGRKEGGFRVYLHOBGVPGSRGSIIVSLPGGDVFGSEFVH 531
 414 IGGWQLAWKWTDK-GEDGKQGGFKRIYHBEVGSASRRGSIIVSPGEG-----EFVQ 466
 532 AALVVSQALFSKGLAEPRMSDAAMVHPSEVAAGSRWKDLPEPGVRRALLVGVGIQILQ 591
 467 AALVVSQALFSKGLAEPRMSDAAMVHPSEVAAGSRWKDLPEPGVRRALLVGVGIQILQ 526
 592 QFAGINGVLYTTPQILEQAGVAVILSKFGLSSASASIISSUTLLMLPCIGFAMLLMDL 651
 527 QFSGINGVLYTTPQILEEAGVELLSDIGIGSESASFLISAFITFLMLPCIGVAMKLMV 586
 652 SGRFLLLOTIPILIASVILVSNLIDGLTALHALLSTVSVIVVFCFCFVMGFGPIPNIL 711
 587 SGRFLLLOTIPILIASVILVSNLIDGLTALHALLSTVSVIVVFCFCFVMGFGPIPNIL 646
 712 CAEIFPVRVGLCIAICAFITWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFF 771
 647 CSEIFPVRVGLCIAICAFITWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFF 706
 772 LKVPETKGMPLVITEFFAVGAKQAAA 798
 707 LKVPETKGMPLVITEFFAVGAKQAAA 733

RESULT 8
 ABU08329 standard; protein; 737 AA.
 ID ABU08329; (first entry)
 AC ABU08329;
 DT 29-MAY-2003 (first entry)
 XX Soybean sugar transport protein #1.
 DE Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
 KW plant sugar transport protein; carbohydrate transport; soybean;
 KW carbohydrate distribution; plant.
 XX Glycine max.
 OS US2002178468-A1.
 FN 28-NOV-2002.
 XX 17-JAN-2002; 2002US-00051902.
 XX 24-APR-1998; 98US-0083044P.
 PR 14-APR-1999; 99US-00291922.
 XX (ALIE/) ALLEN S M.
 PA (HITZ/) HITZ W D.
 PA (KINNEY/) KINNEY A J.
 PA (TINGEY/) TINGEY S V.
 XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;
 XX WPI; 2003-340957/32.
 XX N-PSDB; ABX93201.
 XX Novel plant sugar transport proteins and nucleic acid encoding the
 PT protein useful for producing transgenic plants having altered levels of
 PT sugar transport protein.
 XX Claim 5; Fig 1; 56pp; English.
 XX The present invention relates to the isolation of Arabidopsis thaliana-
 CC like or Beta vulgaris-like sugar transport proteins, and the
 CC polynucleotide sequences encoding them. The plant sugar transport
 CC proteins of the invention have been isolated from corn, rice, soybean,
 CC and wheat. The polypeptides of the invention may be used for altering the
 CC level of expression of a sugar transport protein in a host cell, by

transforming a host cell with a chimeric construct encoding all, or a
 CC portion of the sugar transport protein, in sense or antisense
 CC orientation. Particularly the polypeptides may provide a means to
 CC control carbohydrate transport and distribution in plants. ABU08326-
 CC ABU08333 represent Arabidopsis thaliana-like sugar transport proteins
 XX
 XX Sequence 737 AA;
 Query Match 65.6%; Score 2674; DB 6; Length 737;
 Best Local Similarity 69.9%; Pred. No. 3.8e-250;
 Matches 522; Conservative 92; Mismatches 115; Indels 18; Gaps 8;
 QY 56 MSGAVLVAIVASIGNLQGNWNTAAALVLIKKEFQLONEFTVEGLIVSLGATVIT 115
 Db 1 MKGAVLVAIAASIGNFLQGNWNTAGANGYIKDLALGT--TMRERLVVGNLSGATVIT 58
 QY 116 TFGSLPSDSIGRRPMLILSSITLYPFGSLMLWSNNVYLLARFVDGFGIGLAVTLVPLY 175
 Db 59 TCSGPIADWLGRREPMIISVLYFLGLVLMWSNNVYLLARLLDGFGLAVTLVPEY 118
 QY 176 ISETAPSEIRGLNLTLPQFSGGGMFLSYCMVFGNSLSPSPDWRIMIGLVAIPSPFGL 235
 Db 119 ISETAPSEIRGLNLTLPQFSGGGMFLSYCMVFGNSLSPSPDWRIMIGLVAIPSPFGL 178
 QY 236 TIFYPSPRMLVSKGRVAEAKVLOKLRGDDYSGELSLLEGLGVGGDTSEIYIIGP 295
 Db 179 TIFYPSPRMLVSKGRVLEAKVLOKLRGDDYSGELSLLEGLGVGGDTSEIYIIGP 238
 QY 296 ATEAADLVTGDKKEQITLYGPEGQSWIARPSKIPIMLGSVLSLASHGSMVNSQVPLM 355
 Db 239 ADDVADGHEHATEKDKIRLYGSQAGLSWLSKPVGTQSSIG---LASHGSIINQSMPLM 294
 QY 356 DFIIVTFGSHENMPQAG--GSMRSTLPNFGSMFSTQDQAKNQEOWDENLHRDEEYA 413
 Db 295 DELVTLFGSIHEKLPETGARGSMRSTLPNFGSMFSTQDQAKNQEOWDENLHRDEEYA 354
 QY 414 SDGAGDVEDNLHSPULSRQATGAEKQDIVHHGHSAL-SMRRTLLGEG-GDGVSSTD 471
 Db 355 SDATRGSDDNLHSPISRTISLE-KOLPPPSHSGILGSMRRHSSLMQSGEGGSGTG 413
 QY 472 IGGWQLAWKWSKEGNGRKEGGFRVYLHOBGVPGSRGSIIVSLPGGDVFGSEFVH 531
 Db 414 IGGWQLAWKWTDK-GEDGKQGGFKRIYHBEVGSASRRGSIIVSPGEG-----EFVQ 466
 QY 532 AALVVSQALFSKGLAEPRMSDAAMVHPSEVAAGSRWKDLPEPGVRRALLVGVGIQILQ 591
 Db 467 AALVVSQALFSKGLAEPRMSDAAMVHPSEVAAGSRWKDLPEPGVRRALLVGVGIQILQ 526
 QY 592 QFAGINGVLYTTPQILEQAGVAVILSKFGLSSASASIISSUTLLMLPCIGFAMLLMDL 651
 Db 527 QFSGINGVLYTTPQILEEAGVELLSDIGIGSESASFLISAFITFLMLPCIGVAMKLMV 586
 QY 652 SGRFLLLOTIPILIASVILVSNLIDGLTALHALLSTVSVIVVFCFCFVMGFGPIPNIL 711
 Db 587 SGRFLLLOTIPILIASVILVSNLIDGLTALHALLSTVSVIVVFCFCFVMGFGPIPNIL 646
 QY 712 CAEIFPVRVGLCIAICAFITWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFF 771
 Db 647 CSEIFPVRVGLCIAICAFITWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFF 706
 QY 772 LKVPETKGMPLVITEFFAVGAKQAAA 798
 Db 707 LKVPETKGMPLVITEFFAVGAKQAAA 733

RESULT 9
 ADG47912
 ID ADG47912 standard; protein; 737 AA.
 XX AC ADG47912;
 XX DT 11-MAR-2004 (first entry)
 XX

DE XX Corn Arabidopsis-like sugar transport protein #2.
 KW Arabidopsis-like sugar transport protein;
 KW Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;
 KW corn; plant.
 OS Zea mays.
 XX US2002199217-A1.
 XX 26-DEC-2002.
 XX 17-JAN-2002; 2002US-00051909.
 XX 24-APR-1998; 98US-0083044P.
 XX 14-APR-1999; 99US-00291922.
 XX (HELE/) HELENTJARIIS T G.
 XX Helentjaris TG;
 XX WPI; 2004-040967/04.
 XX N-PSDB; ADG47911.
 XX New isolated polynucleotide encoding a polypeptide having sugar transport protein activity, for producing a transformed plant and for use as probes in physical mapping.
 XX Example 3; SEQ ID NO 8; 71pp; English.
 XX The invention relates to Arabidopsis-like or Beta-vulgaris-like sugar transport proteins and their corresponding nucleic acid sequences. The sequences of the invention are useful to transform a cell. These are also useful to produce a transgenic plant. Probes derived from sequences encoding sugar transport protein may be used for physical mapping. The present sequence is corn arabidopsis-like sugar transport protein.
 XX Sequence 737 AA;
 Query Match 65.6%; Score 2674; DB 8; Length 737;
 Best Local Similarity 69.9%; Pred. No. 3.8e-250;
 Matches 522; Conservative 92; Mismatches 115; Indels 18; Gaps 8;
 QY 56 MSGAVLVAVASIGNLIQGDNDNATIAAAYIKKEPOLQNEPTVEGLIVMSLIGATVIT 115
 DB 1 MKGAVLVAVASIGNLIQGDNDNATIAAGANGYIKKDLALGT--TWERLVGMSLIGATVIT 58
 QY 116 TFGSGLSDSISGRRLMLISLILYFESGLIMLWSPNVYVLLARFVDFGIGLAVTLVPLY 175
 DB 59 TCSGFIADWLGRPMMLISSVLFLGSLVWMLSPNVYVLCARLLDQFGIGLAVTLVPLY 118
 QY 176 ISETAPEIRGLMLNTLPQSGSGMFLSYCMVFGMSLSPSPDRIMGLVLAIPSLPFFGL 235
 DB 119 ISETAPEIRGLMLNTLPQSGSGMFLSYCMVFGMSLSPAPSWRLMLGLVLSIPSLYFAL 178
 QY 236 TIFVLPSPRLVSKGRVAAKVLQKRGKDDVSGELSLILEGLEVGDDTSIEEYIIGP 295
 DB 179 TIFVLPSPRLVSKGRVAAKVLQKRGKDDVSGEMALLVGLGGTDSIEEYIIGP 238
 QY 296 ATEAADLVTDGKEITLYCPBEGOSWIARPSKGPIMLGSLVSLASRHGSMVNOQSVPLM 355
 DB 239 ADDVADGHERATEKDKIRLYGSQAGLSWLSKPTVGSSIG---LASHHGSIIINQSMPLM 294
 QY 356 DPVITLFGSVHENPQAG--GSMRSTLFPNFGSMFVTDHAKNEQWDEENLHRODEEYA 413
 DB 295 DPLVITLFGSHEKLPETGARGSMRSTLFPNFGSMFVTDHAKNEQWDEENLHRODEEYM 354
 QY 414 SDGAGGDEYDNLHSLPLSRGATGAEGKDIVHGHGSGAL--SMRQTLLEG-CDGVSSTD 471
 DB 355 SDATRGSDNDNLHSLPLSRGATGAEGKDIVHGHGSGAL--SMRQTLLEG-CDGVSSTD 471
 QY 472 IGGWQLAWKWTDK-GEQKQGGFKRIYLHEEGVSAARRGSIVIPGEG-----EFVQ 466

DB 414 IGGWQLAWKWTDK-GEQKQGGFKRIYLHEEGVSAARRGSIVIPGEG-----EFVQ 466
 QY 532 AAALVSQSALFSKGLAEPMSDAAMVHSEVAAGSRWKDLFEFGVRALLVGVGIQILQ 591
 DB 467 AAALVSQALYSKELIDGHPVGPAMVHPSETASKGFSWKALLEPGVHALVVGIGIQLQ 526
 QY 592 OFAGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISSITLLMLPCTGFMALINDL 651
 DB 527 QFSGINGVLYTTPQILEEAGVEVLLSDIGIGSESASFLISAFITFLMLPCTGFMALINDL 586
 QY 652 SGRRLFLGLTIPILIASILVLSNLIIDLTGLAHALLSTVSIVVYFCCFVMGFGPIPNIL 711
 DB 587 SGRRLQLLTIPVLIVSLIILVIGSLVNFVNAHAISTVCVVYFCCFVMGFGPIPNIL 646
 QY 712 CABIPFTRVRGLCIAICAPTFWIGDIIIVYSLPVMNAIGLAGVPSIYAVVCLISFVFP 771
 DB 647 CSEIFPTRVRGLCIAICALVFWIGDIIITISLPVNLGSLGLGVFAIYAVVCFISWIEVF 706
 QY 772 LKVPETKGMPLVITEFFAVGAKQAAA 798
 DB 707 LKVPETKGMPLVITEFFAVGAKQAAA 733
 RESULT 10
 ABB93252
 ID ABB93252 standard; protein; 729 AA.
 XX ABB93252;
 XX 31-MAY-2002 (first entry)
 DE Herbicidally active polypeptide SEQ ID NO 2463.
 KW Herbicidal; plant; agriculture; herbicide.
 XX Arabidopsis thaliana.
 XX WO200210210-A2.
 XX 07-FEB-2002.
 XX 28-AUG-2001; 2001WO-EP009892.
 XX 28-AUG-2001; 2001WO-EP009892.
 XX (FARB) BAYER AG.
 XX Tietjen K, Weidner M;
 XX WPI; 2002-269010/31.
 XX Identifying plant target proteins for herbicidally active compounds, comprising aligning and comparing nucleic acid or amino acid sequences from plant with nucleic acid or amino acid sequences from non-plant organisms.
 XX Claim 5; SEQ ID NO 2463; 261pp + Sequence Listing; English.
 XX The invention relates to identifying target proteins (ABB90790-ABB94016) for herbicidally active compounds, comprising aligning and comparing nucleic acid or amino acid sequences from plant with nucleic acid or amino acid sequences from non-plant organisms using suitable search parameters, where plant sequences having an E-value greater by a factor of 3 than the E-value of most similar non-plant sequences are selected. The polypeptides or nucleic acids encoding them are useful for identifying modulators. The identified modulators are useful as herbicides
 XX Sequence 729 AA;
 Query Match 63.4%; Score 2584; DB 5; Length 729;
 Best Local Similarity 68.8%; Pred. No. 2.1e-241;
 Matches 512; Conservative 89; Mismatches 123; Indels 20; Gaps 8;

```
Qy 56 MSGAVLVAIVASIGNLQGDNDATIAAAVLYIKKEFOIQNEPTVEGLIVSMISLIGATIVT 115
D 1 MSGAVLVAIAAAGVNLQGDNDATIAAGAVLYIKKEFNLESNPSEVGLIVAMSLIGATLIT 60
Qy 116 TFSQPLSDSIRGRPMILSSILYFFSGILMWSNVYVLLARFVDGFGIGGLAVTLVPLY 175
D 61 TCSGGVADWLGRPMILSSILYFFSGILMWSNVYVLLARFVDGFGIGGLAVTLVPLY 120
Qy 176 ISETAPSEIRGLNLTLPQFGSGCMFLSYCMVFOMSLSPSPDWIRMLGVLAISLFFPGL 235
D 121 ISETAPSEIRGLNLTLPQFGSGCMFLSYCMVFOMSLSPSPDWIRMLGVLAISLFFPGL 180
Qy 236 TIFVLPSPRWLVSKGRMAEKVLOKRGKDDVSGELSLLEGLVGGTSTIEEYIIGP 295
D 181 TVFFLPSPRWLVSKGRMLEAKRVQLRGREDVSGEVALVVEGLIGIGTIEEYIIGP 240
Qy 296 ATEAADLVTDGDKQITLYGPBEGQSWIARPSKGPIMLGSVLASLRHGS-MVNQSVPL 354
D 241 ADEVTDHDIADVDDQIKLYGAEEGLSWAPVKG---GSTMSVLSRHGSTMSRRQGS 296
Qy 355 MDPIVTLFGSVHENWPOAGGSMRSTLFPNFGSMFSTVDQHAKNEQWDEENLHDDDEYAS 414
D 297 IDPLVTLFGSVHEKMPDT-GSMRSALFPHPFGSMFSGVGNQPRHEDWDEENLVGEGEDYPS 355
Qy 415 DGAGGDVEDNLHSPILSRQATGAEKDIHVHHRGSGALSMMRQTLLEGEGDVSSTDIG 474
D 356 D-HGDDSEDDLHSPILSRQATGAEKDIHVHHRGSGALSMMRQTLLEGEGDVSSTDIG 413
Qy 475 GWQLAWKWSKEGKNGRKEGKRVYHQGVGSGRRGSIIVSLPGGDVFESEFVHAAA 534
D 414 GWQVAKWTEREDESQKE-----EGFPGRSGSIIVSLPGGDGTGE-ADVQASA 462
Qy 535 LVQSALFSKGLAEPRMSDAAMVHPSEVAAGSKRWKDLFEPGVRRLVGVGIQIQOFA 594
D 463 LVSPQALYSKDLLEHTTIGPAMVHPSE-TTKGSIWHDHDFGVKRALVGVGLQIQOFS 521
Qy 595 GINGVLYTTPQILEQAGVAVILSKFGLSSASILISLTLMLPCIGFAMLMDSGR 654
D 522 GINGVLYTTPQILEQAGVAVILSKFGLSSASILISLTLMLPCIGFAMLMDSGR 581
Qy 655 RFLLLGTPILIASLIVVSNLIDGLTAHALISTVSVIVVFCFVMGPGPIENILCAE 714
D 582 RFLLLGTPILIASLIVVSNLIDGLTAHALISTVSVIVVFCFVMGPGPIENILCAE 641
Qy 715 IFPTVRGICIAICFTWIGDIIVTYSPLVNLNAIGLAGVFSIYAVVCLISFVVFVKV 774
D 642 IFPTVRGICIAICFTWIGDIIVTYSPLVNLNAIGLAGVFSIYAVVCLISFVVFVKV 701
Qy 775 PETKGMPLVITEFFFAVGAQAAA 798
D 702 PETKGMPLVITEFFFAVGAQAAA 725

RESULT 11
ADG47933
ID ADG47933 standard; protein; 729 AA.
XX
AC ADG47933;
XX
DT 11-MAR-2004 (first entry)
XX
DE Arabidopsis thaliana-like sugar transport protein #1.
XX
KW Arabidopsis-like sugar transport protein;
KW Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;
KW mouse-ear cress; plant.
XX
OS Arabidopsis thaliana.
XX
PN US2002199217-A1.
XX
PD 26-DEC-2002.
```

```
XX 17-JAN-2002; 2002US-00051909.
XX
XX 24-APR-1998; 98US-0083044P.
XX 14-APR-1999; 99US-00291922.
XX
XX (HELE/) HELENTJARIIS T G.
XX
XX Helentjaris TG;
XX
XX WPI; 2004-040967/04.
XX
XX New isolated polynucleotide encoding a polypeptide having sugar transport
XX protein activity, for producing a transformed plant and for use as probes
XX in physical mapping.
XX
XX Example 3; SEQ ID NO 29; 71pp; English.
XX
XX The invention relates to Arabidopsis-like or Beta-vulgaris-like sugar
XX transport proteins and their corresponding nucleic acid sequences. The
XX sequences of the invention are useful to transform a cell. These are also
XX useful to produce a transgenic plant. Probes derived from sequences
XX encoding sugar transport protein may be used for physical mapping. The
XX present sequence is wheat Arabidopsis thaliana-like sugar transport
XX protein.
XX
XX Sequence 729 AA;
XX
XX Query Match 63.4%; Score 2584; DB 8; Length 729;
XX Best Local Similarity 68.8%; Pred. No. 2.1e-241;
XX Matches 512; Conservative 89; Mismatches 123; Indels 20; Gaps 8;
XX
Qy 56 MSGAVLVAIVASIGNLQGDNDATIAAAVLYIKKEFOIQNEPTVEGLIVSMISLIGATIVT 115
D 1 MSGAVLVAIAAAGVNLQGDNDATIAAGAVLYIKKEFNLESNPSEVGLIVAMSLIGATLIT 60
Qy 116 TFSQPLSDSIRGRPMILSSILYFFSGILMWSNVYVLLARFVDGFGIGGLAVTLVPLY 175
D 61 TCSGGVADWLGRPMILSSILYFFSGILMWSNVYVLLARFVDGFGIGGLAVTLVPLY 120
Qy 176 ISETAPSEIRGLNLTLPQFGSGCMFLSYCMVFOMSLSPSPDWIRMLGVLAISLFFPGL 235
D 121 ISETAPSEIRGLNLTLPQFGSGCMFLSYCMVFOMSLSPSPDWIRMLGVLAISLFFPGL 180
Qy 236 TIFVLPSPRWLVSKGRMAEKVLOKRGKDDVSGELSLLEGLVGGTSTIEEYIIGP 295
D 181 TVFFLPSPRWLVSKGRMLEAKRVQLRGREDVSGEVALVVEGLIGIGTIEEYIIGP 240
Qy 296 ATEAADLVTDGDKQITLYGPBEGQSWIARPSKGPIMLGSVLASLRHGS-MVNQSVPL 354
D 241 ADEVTDHDIADVDDQIKLYGAEEGLSWAPVKG---GSTMSVLSRHGSTMSRRQGS 296
Qy 355 MDPIVTLFGSVHENWPOAGGSMRSTLFPNFGSMFSTVDQHAKNEQWDEENLHDDDEYAS 414
D 297 IDPLVTLFGSVHEKMPDT-GSMRSALFPHPFGSMFSGVGNQPRHEDWDEENLVGEGEDYPS 355
Qy 415 DGAGGDVEDNLHSPILSRQATGAEKDIHVHHRGSGALSMMRQTLLEGEGDVSSTDIG 474
D 356 D-HGDDSEDDLHSPILSRQATGAEKDIHVHHRGSGALSMMRQTLLEGEGDVSSTDIG 413
Qy 475 GWQLAWKWSKEGKNGRKEGKRVYHQGVGSGRRGSIIVSLPGGDVFESEFVHAAA 534
D 414 GWQVAKWTEREDESQKE-----EGFPGRSGSIIVSLPGGDGTGE-ADVQASA 462
Qy 535 LVQSALFSKGLAEPRMSDAAMVHPSEVAAGSKRWKDLFEPGVRRLVGVGIQIQOFA 594
D 463 LVSPQALYSKDLLEHTTIGPAMVHPSE-TTKGSIWHDHDFGVKRALVGVGLQIQOFS 521
Qy 595 GINGVLYTTPQILEQAGVAVILSKFGLSSASILISLTLMLPCIGFAMLMDSGR 654
D 522 GINGVLYTTPQILEQAGVAVILSKFGLSSASILISLTLMLPCIGFAMLMDSGR 581
Qy 655 RFLLLGTPILIASLIVVSNLIDGLTAHALISTVSVIVVFCFVMGPGPIENILCAE 714
```

Db 582 RTLLTTPILASLLVLSLHNSIVHVLSTVSVLYFCFFVNGFGFAPNLCSE 641
 QY 715 IFPTRVRGICIAICAFTEWIGDIIIVTYSIPVLMNAIGLAGVFSIYAVVCLISFVFVFLKV 774
 Db 642 IFPTRVRGICIAICALTWICDIIIVTYSIPVLLKISIGLAGVFGMAIVCCISWVFVFIKV 701
 QY 775 PTKGMPLVITEFPFVAGAKQAAA 798
 Db 702 PETKGMPLVITEFPFVAGAKQAAA 725
 RESULT 12
 ABB91078
 ID ABB91078 standard; protein; 734 AA.
 AC ABB91078;
 XX
 DT 31-MAY-2002 (first entry)
 DE Herbicidally active polypeptide SEQ ID NO 289.
 DE
 XX
 KW Herbicidal; plant; agriculture; herbicide.
 OS Arabidopsis thaliana.
 XX
 PN WO200210210-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 28-AUG-2001; 2001WO-EP009892.
 XX
 PR 28-AUG-2001; 2001WO-EP009892.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Tietjen K, Weidler M;
 XX
 DR WPI; 2002-269010/31.
 XX
 PT Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms.
 XX
 PS Claim 5; SEQ ID NO 289; 261pp + Sequence Listing; English.
 XX
 CC The invention relates to identifying target proteins (ABB90790-ABB94016)
 CC for herbicidally active compounds, comprising aligning and comparing
 CC nucleic acid or amino acid sequences from plant with nucleic acid or
 CC amino acid sequences from non-plant organisms using suitable search
 CC parameters, where plant sequences having an E-value greater by a factor
 CC of 3 than the E-value of most similar non-plant sequences are selected.
 CC The polypeptides or nucleic acids encoding them are useful for
 CC identifying modulators. The identified modulators are useful as
 CC herbicides
 XX
 SQ Sequence 734 AA;
 Query Match 55.8%; Score 2275; DB 5; Length 734;
 Best Local Similarity 61.2%; Pred. No. 2.5e-211;
 Matches 462; Conservative 111; Mismatches 146; Indels 36; Gaps 14;
 QY 56 MSGAVLVAIVASIGNLQGWDNATIAAALVYIKKEPOLQNEPT-VEGLIVMSLIGATIV 114
 Db 1 MKGATLVALAATIGNFLQGWDNATIAAGAMYINKDLNL--PTSVQGLVAVMSLIGATIV 57
 QY 115 TTFSGPLSDSIRGRPMILLSILYFFSGILMLWSPNVYVLLARFVDGFGIGLAVTLVPL 174
 Db 58 TTCGPTSDWLCRRPMLLSVMTYFVCGILMLWSPNVYVLCFARLLNFRAGLAVTLVPL 117
 QY 175 YISIAPESEIRGLNTLPQFSGGGMFLSYCMVFTVNSLSDSPSWRAMLGVLSIPSLLYLF 177

Db 118 YISETAPEIRGLNTLPQFSLGGGMFLSYCMVFTVNSLSDSPSWRAMLGVLSIPSLLYLF 177
 QY 235 LTIFYPESPRMLVSKGRMAEAKVLQKLRGDDVSGELSLLEGLLEVGGDTSIYEYITG 294
 Db 178 LTIFYPESPRMLVSKGRMDEAKVLQKLRGDDVSGELSLLEGLLEVGGDTSIYEYITG 237
 QY 295 PATEAADLVTGDKKE-QITLYGPBEGQSWIARPSKGPIMLGSVLSARHSGSMNQSPV 353
 Db 238 LEDHEGDDTLETVDGQMLYGTENQSYLARPVEQ---NSSLGLSRHSGSLANQSMI 294
 QY 354 LMDPIVTLFGSVHNNPQAGGSMRSTLFPNFGSMFSTVDQ--HAKNEOWD---EENLHRD 408
 Db 295 LKDPVNLVFGSLHERKMPPEAGNTRSGIFPHFGSMFSTTADAPHGKPAHWKEKDIESHYNKD 354
 QY 409 DEYASDGAGDYED---NLHSPLLSRQATGAEGKDIIVHGHGRGALSNNRRTLLGEGGD 465
 Db 355 NDDYATDDGAGDDDDNDLRSLPMSRQTTSD-KDMIPHP7SGTSLSMRRHSTLMQ-GN 412
 QY 466 GVSSTDIGGGMOLAWKWKSEKENGKKEGFKRVYLHOEGVPGSGRRGSTVSLPGGDDVFE 525
 Db 413 GSSMGIGGWHGKRYENDE-----YKRYLKEDGAB-SRRGSIISIPGPD--G 460
 QY 526 GSEFVHAALVSQSALFSKGLAEPMSDAAMVHPSEVAAGSRWKDLFEPGVRRALLVGV 585
 Db 461 GGSYTHASALVSRVSLGPKS-----VHGSAMVPEPEKIAASGSLGALLLEPGVKRALVGV 515
 QY 586 GTIILCOFAGINGVLYYTPQILEQAGVAVILSKFGLSSASAILISLTLMLPCIGFA 645
 Db 516 GTIILCOFAGINGVLYYTPQILERAGVDILSSLSGLSSISASFLISGUTTLMLPAIWA 575
 QY 646 MLLMDLSGRRFLLLGTIPILTASLVLVVSNLIDLTGLAHALLSTVSVIVFCFWMGFG 705
 Db 576 MRLMDVSGRRSLLLTWTIPVLIVSLVIVSELHISKVNAALSTGCVLVFCFWMGFG 635
 QY 706 PIPNLCABIPFTRVRGICIAICAFTEWIGDIIIVTYSIPVLMNAIGLAGVFSIYAVVCL 765
 Db 636 PIPNLCSEIFPTRVRGICIAICAMVFWIGDIIIVTYSIPVLLSSIGLVGVFSIYAAVCV 695
 QY 766 SFVVFVFLKVPETKGMPLVITEFPFVAGAKQAAA 800
 Db 696 SWIFVMKVPETKGMPLVITEFPFVAGAKQAAA 729
 RESULT 13
 ABB92622
 ID ABB92622 standard; protein; 729 AA.
 XX
 AC ABB92622;
 XX
 DT 31-MAY-2002 (first entry)
 DE Herbicidally active polypeptide SEQ ID NO 1833.
 DE
 XX Herbicidal; plant; agriculture; herbicide.
 XX Arabidopsis thaliana.
 OS
 PN WO200210210-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 28-AUG-2001; 2001WO-EP009892.
 XX
 PR 28-AUG-2001; 2001WO-EP009892.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Tietjen K, Weidler M;
 XX
 DR WPI; 2002-269010/31.
 XX
 PT Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences

PT from plant with nucleic acid or amino acid sequences from non-plant
XX organisms.
PS Claim 5; SEQ ID NO 1833; 261pp + Sequence Listing; English.
XX
XX The invention relates to identifying target proteins (ABB90790-ABB94016)
CC for herbicidally active compounds, comprising aligning and comparing
CC nucleic acid or amino acid sequences from plant with nucleic acid or
CC amino acid sequences from non-plant organisms using suitable search
CC parameters, where plant sequences having an E-value greater by a factor
CC of 3 than the E-value of most similar non-plant sequences are selected.
CC The polypeptides or nucleic acids encoding them are useful for
CC identifying modulators. The identified modulators are useful as
CC herbicides
XX
XX Sequence 729 AA;
XX
Query Match 55.1%; Score 2245; DB 5; Length 729;
Best Local Similarity 60.4%; Pred. No. 2.1e-208;
Matches 457; Conservative 110; Mismatches 144; Indels 46; Gaps 14;
QY 56 MSGAVLAIVASIGNLQGNATATAAVLYIKKEFOLENEPTVEGLIVMSLIGATIVT 115
DB 1 MRSVVLVAALAAAGNMLQGNATAGAVIYIKKEFKLEKEPKLEGLIVMSLIGATIT 60
QY 116 TFSGLSDSICRRPMLILSSILYFFSGLIMLWSPNYYVLLARFVDFGIGLAVTLVPLY 175
DB 61 TFSGPVSDKVRSMILSSVLYFLSSIVFWSPNYYVLLFARLLDFGIGLAVTLVPLY 120
QY 176 ISETAPSIIRGLNLTLPFSSGGNFIYCNVFGMSLSPSPDWIRMLGLVLAIPSLFPG 235
DB 121 ISETAPSIIRGLNLTLPFSSGGNFIYCNVFGMSLSPSPDWIRMLGLVLAIPSLFPG 180
QY 236 TIFVLPSPRWLVSKGRMAEKVQLKRLGKDDVSGELSLLEGLVGGDTSEIYIIGP 295
DB 181 AAFPLPSPRWLVSKGRMAEKVQLKRLGKDDVSGELSLLEGLVGGDTSEIYIIGP 240
QY 296 ATEAADDIVTDGKEQITLYGPEQSWIARPSKGPIMLGSVLSASRHGSMVNSQVPLM 355
DB 241 DNEENEGNELPRKQIKLYGPEQSWIARPSKGPIMLGSVLSASRHGSMVNSQVPLM 296
QY 356 DPVTLFGSVHNMEE--CAGGSMRSTLFPNFGSMFVTDQAKNEQWDEENLHDDDEYA 413
DB 297 DPLVTLFGSIHNLSPENNASSRMLFPNFGSMFVTDQAKNEQWDEENLHDDDEYA 349
QY 414 SDGAGDYEDNLHSPLLSRQATGAEKDIYVHGHRSALSRRQTL-LGEGDGVSVSTDI 472
DB 350 ---SSDQDNLNLSPLSPQIT--EPDD--YHQRVTGTMHRRQSSLFMANVGETATSI 401
QY 473 GCGWQLAKWSEKENGK-KGGKRVYLHOE-----GVPSRRGSIYSL-PGGGDV 523
DB 402 GCGWQLAKWSEKENGK-KGGKRVYLHOE-----GVPSRRGSIYSL-PGGGDV 461
QY 524 PEGSEFVHAALVSQALFSKGLASPRMSDAAMVHPSEVAAGKSRWKLPEFVRRALLV 583
DB 462 DVNGYVQAALVSQASMMFGKGTAML-----PKEV-KDGPWRLEKPGVKRALMV 514
QY 584 GVGIGQLQFAGINGVLYTPQILFOAGVAVILSKFLGSSASALISLSTLTLMLPCIG 643
DB 515 GVGIGQLQFAGINGVLYTPQILFOAGVAVILSKFLGSSASALISLSTLTLMLPCIG 573
QY 644 FAWLLMDLSGRFLLGIPILIASIVLVSNLIDGLTAAHMLSTVSVIYFCFVWG 703
DB 574 -----LVMSRLMSTIPILISLVTLVGLSVNLGSGINALISTASVTVYLSCFVWG 626
QY 704 FGPIPNILCAEIPFTRVRLGICAIACFTWIGDIIIVTYSLPVMLNAIGLAGVFSIAYVC 763
DB 627 FGPIPNILCSEIFPFSVRGLCITICATFWICDIIIVTTLPLVWLKSIIGIAGVFGIAYVC 686
QY 764 LISFVFLVFLKVPETKMPLEVIETEFVAVGAKQAAAA 800
DB 687 AVAWFVFLVFLKVPETKMPLEVIETEFVAVGAKQAAAA 723

RESULT 14

AAU97205
ID AAU97205 standard; protein; 486 AA.
XX
XX AAU97205;
AC AC
XX 27-AUG-2002 (first entry)
DT
XX Portion of a soybean sugar transport protein encoded by ss1.pk0022.f1.
DE Soybean; Arabidopsis thaliana-like sugar transport protein;
XX carbohydrate transport; grain filling; annual field crop; plant.
KW
XX Glycine max.
OS
XX US6383776-B1.
DN
XX 07-MAY-2002.
PD
XX 14-APR-1999; 99US-00291922.
PF
XX 24-APR-1998; 98US-0083044P.
PR
XX (DUPO) DU PONT DE NEMOURS & CO E I.
PA
XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;
PI
XX WPI; 2002-453364/48.
XX N-PSDB; ABK51966.
DR
XX New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution.
PS
XX Claim 1; Fig 1; 54pp; English.
XX
XX The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport protein
CC or Beta vulgaris-like sugar transport protein. The polynucleotide
CC sequences are useful for altering the level of sugar transport proteins
CC in plants, i.e. for control of carbohydrate transport and distribution in
CC plant cells, e.g. during grain filling of annual field crops (e.g. corn,
CC rice, soybeans, and wheat), and, for studying carbohydrate flows and
CC sugar transport. The polynucleotide sequences can also be used to isolate
CC cDNA sequences and genes that encode homologues of the new proteins. The
CC present sequence represents a portion of a soybean Arabidopsis thaliana-
CC like sugar transport protein
XX
SQ Sequence 486 AA;

Query Match 41.0%; Score 1670.5; DB 5; Length 486;
Best Local Similarity 68.3%; Pred. No. 8.1e-153;
Matches 333; Conservative 57; Mismatches 84; Indels 15; Gaps 8;
QY 308 DXEQITLYGPEQSWIARPSKGPIMLGSVLSASRHGSMVNSQVPLMDPIVTLFGSVHE 367
DB 5 EKQIKLYGPEQSWIARPSKGPIMLGSVLSASRHGSMVNSQVPLMDPIVTLFGSVHE 59
QY 368 NMQAGSMRSTLFPNFGSMFVTDQAKNEQWDEENLHDDDEYAGAGDYEDNLH 427
DB 60 KLPETG---STLFPNFGSMFVTDQAKNEQWDEENLHDDDEYAGAGDYEDNLH 113
QY 428 PLLSRQATGAEKDIYVHGHRSALSRRQTL-LGEGDGVSVSTDI GCGWQLAKWSEK 486
DB 114 PLISRQTTSLD-KDIPHAHNL-A-SNRQGSLLHNSGEPTGSGTGGGQWLAKWSE 171
QY 487 GENGKRGKRVYLHOE GVGPGVRRGSIYSLPGCGDVFESEFVHAALVSQALFSKGL 546
DB 172 GPDGKKEGFKRILYHQDGGSGRRGSSVSLP-GGDLPTDSEVVOAALVSQALYNE 230
QY 547 AEPRMSDAAMVHPSEVAAGKSRWKLPEFVRRALLVGVIGIQLQFAGINGVLYTPQI 606
DB 231 MRQRPVGFAMIHSEITAKGFSWSDLPFPGVKHALIVGCGKQLQFSGINGVLYTPQI 290

607 LEQAGVAVILSKFGLSSASILSSLTLLMLPCIGFAMLLMDLSGRFLLLTIPILI 666
 QY
 291 LEQAGVGLSLGLGSLSSFLISAVTLLMLPCIAIAMLMDISGRRTLLSTIPVLI 350
 Db
 667 ASLVILVSNLIDIGTLAHLASTVSVIYVCCFVWGFGPIPNILCAEIPTRVGLCIA 726
 QY
 351 AALLILVLGSLVDLSTANASISIVVYFCFFVWGFGPIPNILCAEIPTRVGLCIA 410
 Db
 727 ICAPTFWICDIIIVTSLPVLNAIGLAGVFSIYAVVCLISFVFLKVPETKGMPLVIT 786
 QY
 411 ICALTFWICDIIIVTLPVNLNSVGLAGVFGIYAVVCFIANVFLKVPETKGMPLVII 470
 Db
 787 EPPAVGAKQ 795
 QY
 471 EPPSVGAKQ 479
 Db

RESULT 15

ABU08330
 ID ABU08330 standard; protein; 486 AA.
 AC
 XX
 XX
 DT 29-MAY-2003 (first entry)
 XX
 DE Soybean sugar transport protein #2.
 XX
 KW Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
 KW plant sugar transport protein; carbohydrate transport; soybean;
 KW carbohydrate distribution; plant.
 OS
 XX Glycine max.
 XX
 PN US2002178468-A1.
 XX
 PD 28-NOV-2002.
 XX
 PF 17-JAN-2002; 2002US-00051902.
 XX
 PR 24-APR-1998; 98US-0083044P.
 PR 14-APR-1999; 99US-00291922.
 XX
 PA (ALLE/) ALLEN S M.
 PA (HITZ/) HITZ W D.
 PA (KINN/) KINNEY A J.
 PA (TING/) TINGEY S V.
 XX
 PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
 XX
 XX WPI; 2003-340957/32.
 DR N-PSDB; ABX93202.
 XX

Novel plant sugar transport proteins and nucleic acid encoding the
 PT protein useful for producing transgenic plants having altered levels of
 PT sugar transport protein.
 XX
 PS Claim 5; Fig 1; 56pp; English.
 XX
 CC The present invention relates to the isolation of Arabidopsis thaliana-
 CC like or Beta vulgaris-like sugar transport proteins, and the
 CC polynucleotide sequences encoding them. The plant sugar transport
 CC proteins of the invention have been isolated from corn, rice, soybean,
 CC and wheat. The polypeptides of the invention may be used for altering the
 CC level of expression of a sugar transport protein in a host cell, by
 CC transforming a host cell with a chimeric construct encoding all, or a
 CC portion of the sugar transport protein, in sense or antisense
 CC orientation. Particularly the polypeptides may provide a means to
 CC control carbohydrate transport and distribution in plants. ABU08326-
 CC ABU08333 represent Arabidopsis thaliana-like sugar transport proteins
 XX
 SQ Sequence 486 AA;

Query Match 41.0%; Score 1670.5; DB 6; Length 486;
 Best Local Similarity 68.1%; Pred. No. 8.1e-153;
 Matches 333; Conservative 57; Mismatches 84; Indels 15; Gaps 8;
 QY 308 DKEQITLYGPEGQSWIAPRSPKPIMLGSLVSLASRHSVMNQSVPLMDPIVTLFGSVHE 367
 Db 5 EKQIKLYGPEGQSWIAPRSPKPIMLGSLVSLASRHSVMNQSVPLMDPIVTLFGSVHE 59
 QY 368 NMPQAGSMRSLTFNFGSMFVTDQHAQNEQWDEENLHRDDEYASDAGGDDYEDNLHS 427
 Db 60 KLPETG---STLPHFGSMFVGNQPRNEDWDESLAREGDDVSD--AGSDNQLQS 113
 QY 428 PILSRCATCAEGKDIIVHCHRGSAISMRQTLI--GEGDGVSSDTIGGQWLAKWSEKE 486
 Db 114 PLISRQTTSLD-KDIPPHAHSNLA-SMRQGSLLHNSGEPTGTGTGGGWQLAKWSEKE 171
 QY 487 GENGKKEGFKRYLHQEGVPGSRGSIIVSLPGGDFEGSEFVHAAALVSQSALFSKGL 546
 Db 172 GPDGKKEGFKRYLHQEGVPGSRGSIIVSLPGGDFEGSEFVHAAALVSQALYNSDL 230
 QY 547 AEPRMSDAMWHPSEVAAGSRWKDLFERGVTRALLVGVGIILOQFAGINGVLYTTPQI 606
 Db 231 MRQRPVGPAMIHPSETIAKGPSWDLFEPGVKHALIVGVGMQILOQFSGINGVLYTTPQI 290
 QY 607 LEQAGVAVILSKFGLSSASILSSLTLLMLPCIGFAMLLMDLSGRFLLLTIPILI 666
 Db 291 LEQAGVGLSLGLGSLSSFLISAVTLLMLPCIAIAMLMDISGRRTLLSTIPVLI 350
 QY 667 ASLVILVSNLIDIGTLAHLASTVSVIYVCCFVWGFGPIPNILCAEIPTRVGLCIA 726
 Db 351 AALLILVLGSLVDLSTANASISIVVYFCFFVWGFGPIPNILCAEIPTRVGLCIA 410
 QY 727 ICAPTFWICDIIIVTSLPVLNAIGLAGVFSIYAVVCLISFVFLKVPETKGMPLVIT 786
 Db 411 ICALTFWICDIIIVTLPVNLNSVGLAGVFGIYAVVCFIANVFLKVPETKGMPLVII 470
 QY 787 EPPAVGAKQ 795
 Db 471 EPPSVGAKQ 479

Search completed: October 13, 2004, 11:38:26
 Job time : 85.4139 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 11:33:23 ; Search time 24.8337 Seconds
(without alignments)
2136.387 Million cell updates/sec

Title: US-10-051-909-32

Perfect score: 4075

Sequence: 1 IRSSGSLAVQTFPTDLDR.....PLEVITEFFAVGAKQAAXA 800

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3517	86.3	747	3	US-09-291-922-2
2	2674	65.6	737	3	US-09-291-922-8
3	2584	63.4	729	3	US-09-291-922-29
4	1670.5	41.0	486	3	US-09-291-922-10
5	1424	34.9	345	3	US-09-291-922-14
6	943	23.1	228	3	US-09-291-922-16
7	562	13.8	131	3	US-09-291-922-6
8	548	13.4	488	4	US-10-162-012-46
9	496	12.2	523	3	US-09-291-922-24
10	489.5	12.0	513	3	US-09-291-922-20
11	489	12.0	529	3	US-09-291-922-28
12	488	12.0	549	3	US-09-291-922-30
13	485	11.9	510	3	US-09-291-922-22
14	463.5	11.4	539	3	US-09-291-922-26
15	447	11.0	476	4	US-09-489-039A-11933
16	440.5	10.8	514	4	US-09-489-039A-11902
17	427	10.5	501	4	US-09-489-039A-11731
18	426	10.5	521	4	US-09-489-039A-9549
19	421.5	10.3	584	2	US-08-928-692-13
20	421.5	10.3	584	3	US-09-339-972-13
21	384.5	9.4	534	3	US-09-031-392-4
22	384.5	9.4	534	3	US-09-299-549-4
23	384.5	9.4	534	2	US-09-610-417-4
24	365	9.0	493	2	US-09-031-392-10
25	365	9.0	493	3	US-09-299-549-10
26	365	9.0	493	3	US-09-610-417-10
27	358	8.8	488	2	US-08-928-692-10

28 358 8.8 488 3 US-09-339-972-10 Sequence 10, Appli
29 346 8.5 524 2 US-08-928-692-12 Sequence 12, Appli
30 346 8.5 524 3 US-09-339-972-12 Sequence 12, Appli
31 344 8.4 509 2 US-09-031-392-6 Sequence 6, Appli
32 344 8.4 509 3 US-09-299-549-6 Sequence 6, Appli
33 344 8.4 509 3 US-09-610-417-6 Sequence 6, Appli
34 341.5 8.4 504 4 US-09-679-686B-21 Sequence 21, Appli
35 340 8.3 511 4 US-09-679-686B-12 Sequence 12, Appli
36 331 8.1 492 2 US-08-355-844-3 Sequence 3, Appli
37 331 8.1 492 5 PCT-US95-16136-3 Sequence 3, Appli
38 330 8.1 488 2 US-08-928-692-11 Sequence 11, Appli
39 330 8.1 488 3 US-09-339-972-11 Sequence 11, Appli
40 328.5 8.1 863 4 US-09-894-927B-9 Sequence 9, Appli
41 322 7.9 552 4 US-09-248-796A-20713 Sequence 20713, A
42 315.5 7.7 514 4 US-09-679-686B-22 Sequence 22, Appli
43 312 7.7 502 4 US-09-679-686B-2 Sequence 2, Appli
44 310 7.6 494 2 US-09-031-392-5 Sequence 5, Appli
45 310 7.6 494 3 US-09-299-549-5 Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-291-922-2
; Sequence 2, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tirgey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 747
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (129)
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (133).. (134)
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (144)
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (178)
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (207)
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (218)
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (220)
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (236)
US-09-291-922-2

Query Match 86.3%; Score 3517; DB 3; Length 747;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 684; Conservative 27; Mismatches 34; Indels 2; Gaps 2;

```

QY 56 MSGAVLVAVISAGNLLQGWDNATIAAAVLYIKKEFQLONEPTVEGLIVMSLIGATVIT 115
DB 1 MCGAVVVAIAASIGNLLQGWDNATIAAGVLYIKKEFNQSEPLIEGLIVAMFLIGATVIT 60
QY 116 TFSGLPSISGRPMILILSSILYFSGILMLWSNVVYLLARFVDGFGIGLAVTLVPLY 175
DB 61 TSPGRADCVGRPMVAVASVLYFVSGVLWMAPIVYLLARLIDGFGIGLAVTLVPLY 120
QY 176 ISEIAPSEIR-GLNLTLPQFSG-SGGMFLSYCMVFGMSLSPSPDWRLMGLVLAIPSLFF 233
DB 121 ISETAPRXSWCXNLTLPQFIVGVGMFLSYCMVFGMSLMPKPDWRMLMGLVLSPLXYF 180
QY 234 GLTIYLPESPRWLVSKGRMAEAKVLOKLRGDDVSGELSLLEGLEVGGDTSEIEYII 293
DB 181 GLTVFYPESPRWLVSKGRMAEAKVXQRLRGREDSVSEXALLVEGLVGVGKOTRIEXYII 240
QY 294 GPATEAADDLVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSLASRHGSMVNSVP 353
DB 241 GPATEAADDLVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSLASRHGSMVNSVP 300
QY 354 LMDPIVTLFGSVHENMPQAGSMRSTLFPNFGSMFVTDQAKNEQWDEENLHRDDEYA 413
DB 301 LMDPIVTLFGSVHENMPQAGSMRSTLFPNFGSMFVTDQAKNEQWDEENLHRDDEYA 360
QY 414 SDGAGDYEDNLHSPLLSRQATGAEGKDI VHHGRGSGALSMRROTLLGEGDGVSSDID 473
DB 361 SDGAGDYEDNLHSPLLSRQATGAEGKDI VHHGRGSGALSMRROTLLGEGDGVSSDID 420
QY 474 GGWLAWKWSKEGNGRKEGGFKRYVYLHQEGVPGSRGSIIVSLPGGDFVSGSEFVHAA 533
DB 421 GGWLAWKWSKEGNGRKEGGFKRYVYLHQEGVPGSRGSIIVSLPGGDFVSGSEFVHAA 480
QY 534 ALVQSALFSGKLAEPMSDAAMVHPSEVAAGSRWKDLFEPGVRRALLVGVGIQILQQF 593
DB 481 ALVQSALFSGKLAEPMSDAAMVHPSEVAAGSRWKDLFEPGVRRALLVGVGIQILQQF 540
QY 594 AGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLLMLPCIGFAMLLMDLSG 653
DB 541 AGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLLMLPCIGFAMLLMDLSG 600
QY 654 RFLGLGPIPIILASVILVNSLIDGLTLAHLISVIVSVIVYVCCFWMGFGPIPNILCA 713
DB 601 RFLGLGPIPIILASVILVNSLIDGLTLAHLISVIVSVIVYVCCFWMGFGPIPNILCA 660
QY 714 BIFPVRVGLCIAICAFTFWIGDIIVTVSLPVMNAIGLAGVFSIYAVVCLISFVFVLK 773
DB 661 BIFPVRVGLCIAICAFTFWIGDIIVTVSLPVMNAIGLAGVFSIYAVVCLISFVFVLK 720
QY 774 VPETKGMPLVITEFFAVGAKQAQAA 800
DB 721 VPETKGMPLVITEFFAVGAKQAQAA 747

```

RESULT 2

```

US-09-291-922-8
; Sequence 5, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291.922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 737

```

```

; TYPE: PRT
; ORGANISM: Glycine max
; US-09-291-922-8

Query Match 65.6%; Score 2674; DB 3; Length 737;
Best Local Similarity 69.9%; Pred. No. 4.7e-256;
Matches 522; Conservative 92; Mismatches 115; Indels 18; Gaps 8;

QY 56 MSGAVLVAVISAGNLLQGWDNATIAAAVLYIKKEFQLONEPTVEGLIVMSLIGATVIT 115
DB 1 MCGAVVVAIAASIGNLLQGWDNATIAAGVLYIKKEFNQSEPLIEGLIVAMFLIGATVIT 58
QY 116 TFSGLPSISGRPMILILSSILYFSGILMLWSNVVYLLARFVDGFGIGLAVTLVPLY 175
DB 59 TCSGPIADWLGRPRMIISSVLYFLGGLVMLSPNVYVLCARLDGFGIGLAVTLVPLY 118
QY 176 ISEIAPSEIRGLNLTLPQFSGGMFLSYCMVFGMSLSPSPDWRLMGLVLAIPSLFFGL 235
DB 119 ISETAPSEIRGLNLTLPQFSGGMFLSYCMVFGMSLSPSPDWRLMGLVLAIPSLFFGL 178
QY 236 TIFYPESPRWLVSKGRMAEAKVLOKLRGDDVSGELSLLEGLEVGGDTSEIEYIIGP 295
DB 179 TIFYPESPRWLVSKGRMAEAKVLOKLRGDDVSGELSLLEGLEVGGDTSEIEYIIGP 238
QY 296 ATEAADDLVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSLASRHGSMVNSVPLM 355
DB 239 ADDVADGHEHATEKOKIRLYGSAQGLSMLSKPVTQSSIG---LASHHGSIINQSMPLM 294
QY 356 DIPIVTLFGSVHENMPQAG--GSMRSTLFPNFGSMFVTDQAKNEQWDEENLHRDDEYA 413
DB 295 DPLVTLFGSVHENMPQAG--GSMRSTLFPNFGSMFVTDQAKNEQWDEENLHRDDEYM 354
QY 414 SDGAGDYEDNLHSPLLSRQATGAEGKDI VHHGRGSGAL-SMRROTLLGEG-DGVSSSTD 471
DB 355 SDATRGSDDDNLHSPILSRQITSLK-KDLPPTPSHGSLGSMRHSLSMQSGGSGTG 413
QY 472 IGGHQLAWKWSKEGNGRKEGGFKRYVYLHQEGVPGSRGSIIVSLPGGDFVSGSEFVH 531
DB 414 IGGHQLAWKWTDK-GEDEKQGGFKRYLHHEGVASRSGSIIVSIPGEG-----EFVQ 466
QY 532 AAALVSQALFSGKLAEPMSDAAMVHPSEVAAGSRWKDLFEPGVRRALLVGVGIQILQ 591
DB 467 AAALVSQALFSGKLAEPMSDAAMVHPSEVAAGSRWKDLFEPGVRRALLVGVGIQILQ 526
QY 592 QFAGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLLMLPCIGFAMLLMDL 651
DB 527 QFSGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLLMLPCIGFAMLLMDV 586
QY 652 SCRRFLGLGPIPIILASVILVNSLIDGLTLAHLISVIVSVIVYVCCFWMGFGPIPNIL 711
DB 587 SCRRFLGLGPIPIILASVILVNSLIDGLTLAHLISVIVSVIVYVCCFWMGFGPIPNIL 646
QY 712 CAEIPFTRVGLCIAICAFTFWIGDIIVTVSLPVMNAIGLAGVFSIYAVVCLISFVFV 771
DB 647 CSEIPFTRVGLCIAICAFTFWIGDIIVTVSLPVMNAIGLAGVFSIYAVVCLISFVFV 706

QY 772 LKVPETKGMPLVITEFFAVGAKQAQAA 798
DB 707 LKVPETKGMPLVITEFFAVGAKQAQAA 733

```

RESULT 3

```

US-09-291-922-29
; Sequence 29, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291.922

```

; CURRENT FILING DATE: 1999-04-14
 ; EARLIER APPLICATION NUMBER: 60/083,044
 ; EARLIER FILING DATE: April 24, 1998
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 29
 ; LENGTH: 729
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 US-09-291-922-29

Query Match 63.4%; Score 2584; DB 3; Length 729;
 Best Local Similarity 68.8%; Pred. No. 4.1e-247;
 Matches 512; Conservative 89; Mismatches 123; Indels 20; Gaps 8;

Qy	56	MSGAVLVAVIASQNLQGDNDNATIAAVALYIKKEFOLQNEPTVEGLIVMSKIGATIVT	115
Db	1	MSGAVLVAVIAAAGNLLQGDNDNATIAAVALYIKKEFNLESNPVEGLIVAMSLIGATLT	60
Qy	116	TFSGPLSDSIRRRPMLILSSLYFFSLGLIMLSNVVLLARFVDFGFGIAGLVTVLPV	175
Db	61	TCSGGVADWLGRRPMLILSSLYFVGLVLMVSNVTVLLGRLLDGFGVGLVTVLPV	120
Qy	176	ISETAPSEIRGLNLTLPQSGSGMFLSYCMVFGMSLSPDPWRIMLGVAIPSLPFFGL	235
Db	121	ISETAPSEIRGLNLTLPQSGSGMFLSYCMVFGMSLSPDPWRIMLGVAIPSLPFFFL	180
Qy	236	TIFVLPSPRLVSKGMAZAKVLOKLRGDDVSGELSLLEGLVGGDTSTEEYLIIGP	295
Db	181	TVFVLPSPRLVSKGRLMELAKVLOKLRGDDVSGELSLLEGLVGGDTSTEEYLIIGP	240
Qy	296	ATEAADLVTDGDEQITLYPPEGQSWIARPSKXPGMLGSLVSLASRHGS-MVNSVPL	354
Db	241	ADEVTDHDIADVQDKLYGAEEGLSWARPVKG---GSTXSVLSRHGSTMSRRQSL	296
Qy	355	MDPIVTLFGSVHNPQAGGSMRTLPNFGSMFVSVDQAKNQWDEENLHRDDEYAS	414
Db	297	IDPLVTLFGSVHNPQAGGSMRTLPNFGSMFVSVDQAKNQWDEENLHRDDEYAS	355
Qy	415	DGAGGVEDNLHSPULGRQATGAGKQDIVHGHRSALSMRQTLLEGEGDGVSTSDIGG	474
Db	356	D-HGDDSEDDHSPISLSEQTSMR-KDMPHTAHGTLSTFRHGSQVQAGGAGSMGIGG	413
Qy	475	GWQAWKSEKEGNGRKEGFKVYHLEQGVPGSRGSIIVSLPGGQGVFEGSFVHAAA	534
Db	414	GWQAWKSEKEGNGRKEGFKVYHLEQGVPGSRGSIIVSLPGGQGVFEGSFVHAAA	462
Qy	535	LVSQALPESKGLAEPFMSDAAMVHPSEVAAGSRWKDLFEGVRRALLVGVGIIQOQFA	594
Db	463	LVSQALPESKGLAEPFMSDAAMVHPSEVAAGSRWKDLFEGVRRALLVGVGIIQOQFA	521
Qy	595	GINGVLYYTPQILEQAGVAVILSKFGLSSASASILISLTLMLPCIGFAMLLMDLSGR	654
Db	522	GINGVLYYTPQILEQAGVAVILSKFGLSSASASILISLTLMLPCIGFAMLLMDLSGR	581
Qy	655	RELLLTGTPILLASVILVSNLIDLGLTAHALLSTVSIVVFCFVWGFGPIPNILCAE	714
Db	582	RELLLTGTPILLASVILVSNLIDLGLTAHALLSTVSIVVFCFVWGFGPIPNILCAE	641
Qy	715	IFPFRVGLCAICAFTEWIGDIIIVTSLPVMNAIGLAGVFSYAVVCLISFVFLKV	774
Db	642	IFPFRVGLCAICAFTEWIGDIIIVTSLPVMNAIGLAGVFSYAVVCLISFVFLKV	701
Qy	775	PETKGMPLVETFEFFAVGAKQAAA 798	
Db	702	PETKGMPLVETFEFFAVGAKQAAA 725	

RESULT 4
 US-09-291-922-10
 ; Sequence 10, Application US/09291922
 ; Patent No. 6383776
 ; GENERAL INFORMATION:

; APPLICANT: Allen, Steve
 ; APPLICANT: Hitz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB-1163
 ; CURRENT APPLICATION NUMBER: US/09/291,922
 ; CURRENT FILING DATE: 1999-04-14
 ; EARLIER APPLICATION NUMBER: 60/083,044
 ; EARLIER FILING DATE: April 24, 1998
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 10
 ; LENGTH: 486
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 US-09-291-922-10

Query Match 41.0%; Score 1670.5; DB 3; Length 486;
 Best Local Similarity 68.1%; Pred. No. 1.1e-156;
 Matches 333; Conservative 57; Mismatches 84; Indels 15; Gaps 8;

Qy	308	DKEQITLYPPEGQSWIARPSKXPGMLGSLVSLASRHGSVWNCQVPLMDPIVTLFGSVH	367
Db	5	EKDQIKLYPPEGQSWIARPSKXPGMLGSLVSLASRHGSVWNCQVPLMDPIVTLFGSVH	59
Qy	368	NMPQAGGSMRSTLPNFGSMFVSVDQAKNQWDEENLHRDDEYASDAGGAGDYEDNLHS	427
Db	60	KLPETG---STLFPFGSMFVSVDQAKNQWDEENLHRDDEYASDAGGAGDYEDNLHS	113
Qy	428	PILSRQATCAEKKDVIHGHRSALSMRQTL-GEQGDVSVSTDIQGGWQAWKSEKE	486
Db	114	PLISRQATCAEKKDVIHGHRSALSMRQTL-GEQGDVSVSTDIQGGWQAWKSEKE	171
Qy	487	GENGRKEGFKVYHLEQGVPGSRGSIIVSLPGGQGVFEGSFVHAAALVQSALFSKGL	546
Db	172	GDGKKEGFKVYHLEQGVPGSRGSIIVSLPGGQGVFEGSFVHAAALVQSALFSKGL	230
Qy	547	ASERMSDAAMVHPSEVAAGSRWKDLFEGVRRALLVGVGIIQOQFAGINGVLYYTPQI	606
Db	231	MRQRPVGPAMIHPSSETIAKPSWSDLFEGVRRALLVGVGIIQOQFAGINGVLYYTPQI	290
Qy	607	LEQAGVAVILSKFGLSSASASILISLTLMLPCIGFAMLLMDLSGRFLLIGTIPILI	666
Db	291	LEQAGVAVILSKFGLSSASASILISLTLMLPCIGFAMLLMDLSGRFLLIGTIPILI	350
Qy	667	ASLIVLVSNLIDLGLTAHALLSTVSIVVFCFVWGFGPIPNILCAEIFPFRVGLCIA	726
Db	351	AALLILVGLVDLGSSTANASISTISVIVFCFVWGFGPIPNILCAEIFPFRVGLCIA	410
Qy	727	ICAFTEWIGDIIIVTSLPVMNAIGLAGVFSYAVVCLISFVFLKVETKGMPLVET	786
Db	411	ICAFTEWIGDIIIVTSLPVMNAIGLAGVFSYAVVCLISFVFLKVETKGMPLVET	470
Qy	787	EFFAVGAKQ 795	
Db	471	EFFAVGAKQ 479	

RESULT 5
 US-09-291-922-14
 ; Sequence 14, Application US/09291922
 ; Patent No. 6383776
 ; GENERAL INFORMATION:

; APPLICANT: Allen, Steve
 ; APPLICANT: Hitz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB-1163
 ; CURRENT APPLICATION NUMBER: US/09/291,922
 ; CURRENT FILING DATE: 1999-04-14
 ; EARLIER APPLICATION NUMBER: 60/083,044

; EARLIER FILING DATE: April 24, 1998
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 14
 ; LENGTH: 345
 ; TYPE: PRT
 ; ORGANISM: Triticum aestivum
 US-09-291-922-14

Query Match 34.9%; Score 1424; DB 3; Length 345;
 Best Local Similarity 82.9%; Pred. No. 1.9e-132;
 Matches 281; Conservative 24; Mismatches 32; Indels 2; Gaps 2;

QY 462 EGGDVSSTDTGGWQLAWKSEKGEGRKGGFKVYLHOEGVPGSRGSIIVSLPGGG 521
 Db 4 EGGAVSTGTGGWQLAWKSEKGEGRKGGFKVYLHOEGVADSRGSGVVSPLPGGG 63
 QY 522 DVFE-GSEFVHAALVSOSALFSGKLAEPRM-SDAAMVHPSEVAAGSRWKDLFEPGVRR 579
 Db 64 DATQGGSGFIHAAALVSHSALYSKDLMEERMAAGPAMHPLEAAPKGSINWDLFEPGVRR 123
 QY 580 ALLVGVGTIILQOPAGVNGVLYTPQILEQAGVAVILSKGLSSASASILISLTLML 639
 Db 124 ALFVGVGTIILQOPAGVNGVLYTPQILEQAGVAVILSKGLSSASASILISLTLML 183
 QY 640 PCIGFAMLLDLSGRFFLLGTFIPILIASIVLVSNLIDGLTLAHALLSTVSIVVYFCC 699
 Db 184 PSIGVAMELMDISGRFFLLGTFIPILIASIVLVSNLIDGLTLAHALLSTVSIVVYFCC 243
 QY 700 FVMGFGPIPNILCAEIPFTRVRGLCIACTFWIGDIIIVTSLPVMNLGAGVPSIY 759
 Db 244 FVMGFGPIPNILCAEIPFTRVRGVCIACTFWIGDIIIVTSLPVMNLGAGVPSIY 303
 QY 760 AVVCLISFVFLKVPETKGMPLVTEFFFAVGAQAQA 798
 Db 304 AVVCLIAFVFLKVPETKGMPLVTEFFFAVGAQAQA 342

RESULT 6

US-09-291-922-16
 ; Sequence 16, Application US/09291922
 ; Patent No. 6383776
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Hitz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB-1163
 ; CURRENT APPLICATION NUMBER: US/09/291,922
 ; CURRENT FILING DATE: 1999-04-14
 ; EARLIER APPLICATION NUMBER: 60/083,044
 ; EARLIER FILING DATE: April 24, 1998
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 16
 ; LENGTH: 228
 ; TYPE: PRT
 ; ORGANISM: Triticum aestivum
 US-09-291-922-16

Query Match 23.1%; Score 943; DB 3; Length 228;
 Best Local Similarity 80.6%; Pred. No. 5.4e-85;
 Matches 179; Conservative 25; Mismatches 18; Indels 0; Gaps 0;

QY 574 EPGVRALLVGVGTIILQOPAGVNGVLYTPQILEQAGVAVILSKGLSSASASILISL 633
 Db 1 EPGVKHALFVGIQILQOPAGVNGVLYTPQILEQAGVAVILSKGLSSASASILISL 60
 QY 634 TTLMLPCIGFAMLLDLSGRFFLLGTFIPILIASIVLVSNLIDGLTLAHALLSTVS 693
 Db 61 TTLMLPSIGIAMLMDSGRFFLLSTIPILIALVVLVNVLDVGNVHAALSTISV 120

QY 694 IYVFCFVWGFGPIPNILCAEIPFTRVRGLCIACTFWIGDIIIVTSLPVMNLGAG 753
 Db 121 IYVFCFVWGFGPIPNILCAEIPFTRVRGICIACTFWIGDIIIVTSLPVMNLGAG 180
 QY 754 GVFSIYAVVCLISFVFLKVPETKGMPLVTEFFFAVGAQA 795
 Db 181 GVFGIYAVVCLAFVFMVYKVPETKGMPLVTEFFFAVGAQA 222

RESULT 7

US-09-291-922-6
 ; Sequence 6, Application US/09291922
 ; Patent No. 6383776
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Hitz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB-1163
 ; CURRENT APPLICATION NUMBER: US/09/291,922
 ; CURRENT FILING DATE: 1999-04-14
 ; EARLIER APPLICATION NUMBER: 60/083,044
 ; EARLIER FILING DATE: April 24, 1998
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 6
 ; LENGTH: 131
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 US-09-291-922-6

Query Match 13.8%; Score 562; DB 3; Length 131;
 Best Local Similarity 80.5%; Pred. No. 1.4e-47;
 Matches 103; Conservative 17; Mismatches 8; Indels 0; Gaps 0;

QY 669 LVILVSNLIDGLTLAHALLSTVSIVVYFCCFVMGFGPIPNILCAEIPFTRVRGLCIA 728
 Db 1 VLTILVNLIDVGTVMVHASLSTVSIVVYFCCFVMGFGPIPNILCAEIPFTRVRGLCIA 60
 QY 729 AFTFWIGDIIIVTSLPVMNLGAGVPSIYAVVCLISFVFLKVPETKGMPLVTEF 788
 Db 61 ALTFWIGDIIIVTSLPVMNLGAGVPSIYAVVCLISFVFLKVPETKGMPLVTEF 120
 QY 789 FAVGAQA 796
 Db 121 FSVGAQA 128

RESULT 8

US-10-162-012-46
 ; Sequence 46, Application US/10162012
 ; Patent No. 6682597
 ; GENERAL INFORMATION:
 ; APPLICANT: Curtis, Rory A.J.
 ; APPLICANT: Silos-Santiago, Inmaculada
 ; APPLICANT: Gu, Wei
 ; TITLE OF INVENTION: NOVEL HUMAN ION CHANNEL AND TRANSPORTER FAMILY MEMBERS
 ; FILE REFERENCE: 10448-190001
 ; CURRENT APPLICATION NUMBER: US/10/162,012
 ; CURRENT FILING DATE: 2002-06-04
 ; PRIOR APPLICATION NUMBER: US 60/209,845
 ; PRIOR FILING DATE: 2000-06-06
 ; PRIOR APPLICATION NUMBER: US 09/875,321
 ; PRIOR FILING DATE: 2001-06-06
 ; PRIOR APPLICATION NUMBER: PCT/US01/18340
 ; PRIOR FILING DATE: 2001-06-06
 ; PRIOR APPLICATION NUMBER: US 60/209,257
 ; PRIOR FILING DATE: 2000-06-05
 ; PRIOR APPLICATION NUMBER: US 09/875,423
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: PCT/US01/18398
 ; PRIOR FILING DATE: 2001-06-05

```

; PRIOR APPLICATION NUMBER: US 60/209,238
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/875,363
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US01/18247
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/227,068
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 09/928,530
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: PCT/US01/25475
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: US 60/226,770
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/934,421
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26096
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/279,281
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 10/109,029
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: PCT/US02/09728
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/290,288
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US (not assigned)
; PRIOR FILING DATE: 2002-05-13
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-10-162-012-46

Query Match      13.4%; Score 548; DB 4; Length 488;
Best Local Similarity 22.1%; Pred. No. 3.6e-45;
Matches 171; Conservative 105; Mismatches 167; Indels 330; Gaps 16;

QY 62 VAIVASI-GNLLQGDWNTAAAVLYIKKEF-----QLONEPTV-EGLIVSMS 107
DB 1 VALVAALGGGFLGVDYTGVIIGFLALIDFLFRFGLTSSGALAEVLGVSTVLTGLVVISIF 60
QY 108 LIGATIVTFSGPLSDSGRRPMLLSILYFFSGL-----IWLKSPNVYVLLA 157
DB 61 FLGRLLGSLFACKLGRFRKXKSLTIALVLFVIGALLSGAAPGVTTIGLWA--FYLLIVG 118
QY 158 RFVDGFGIGLAVTLVPLVYIETAPSEIRGLNLTLPQFSGSGMFLSYCMWFGMSLSPSPD 217
DB 119 RVLVGLGVGGASVLPVMIYSEIAPKALRGALGSLYQLAITIGILVA--AIIGLGKNTNN 176
QY 218 -----WRIMGLVLAIPSLFFFGLTITVYLPESPRMLVSGRMAEAKVLOKLRGKDDV 269
DB 177 DSALNSWGRIPGLQLVPALLL--IGLFTLPESPRMLVSRKLEAREVLAKLRGVEDV 235
QY 270 SGEILLSLLEGLVGGDTSEIEYIIGPATEADDLVTDGDKQITLYGPEEQSWIARPSK 329
DB 236 DOEIEIKAELEA----- 248
QY 330 GFIMLGSVLASRHGSMVNSQVPLMDPIVTLFGSVHENMPQAGSMESTLFPNFGMFS 389
DB 249 -----T 249
QY 390 VTDQAKNEQWDEENLHRDDEYASDGAGDYEDNLHSPLLSRQATGAEGKDIVHGHGRG 449
DB 250 VSEENKAGKASW----- 260
QY 450 SALSMMRQTLLEGCGDGVSSTDIGGQWLAWKWEKEGNGRKEGGFKRVYVHOBGVPGS 509
DB 261 ----- 260
```

```

QY 510 RRGSTVSLPGGDVFEQSEFVHAALVSQALFSKGLAEPRMSDAAMVHPSEVAAGSRW 569
DB 261 -----GELFRGET----- 268
QY 570 KDLFEPGVRALLVGVGHILOQFAGINGVLYYTPQILEQAGVAVILSKFGISSASAIL 629
DB 269 ----RPKVRQRLLMGVMQLQAQQLTGINAIFYSTIFKSVGV-----SDSVASIL 315
QY 630 ISSLTLLMLPCIGFAML-LMDLSORR-FLLLGTTPILLIASLV-----ILVYSNLID 679
DB 316 VTIIIGVWNVFTFVALIFLVDRFGRRELLLLGAAGMAICFLILGASIGVALLLNKPKD 375
QY 680 LCTLAHALLSTVSVIVYFCCFUMGFGPIPNILCABIFPTRVRGCLCIAICAFTFWIGDIIV 739
DB 376 PSSKAAGIVAIIVFILLFIAFFALGAGPIFWILSELFTKVRKSKALATAANWLANFII 435
QY 740 TYSLPVMLNAIGLA---GVFSIYAVVCLISFVVFVFLKVPETKGMPLVITEFF 789
DB 436 GFLFPYITGAIGLALGGVYVFLVFAGLLVFLFVFFVFPETKGRTEIEELF 488

RESULT 9
US-09-291-922-24
; Sequence 24, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 24
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Glycine max
US-09-291-922-24

Query Match      12.2%; Score 496; DB 3; Length 523;
Best Local Similarity 21.9%; Pred. No. 5.9e-40;
Matches 165; Conservative 109; Mismatches 189; Indels 290; Gaps 18;

QY 59 AVLVAVASIGNLLQGDWNTAAAVLYIKKEFQLONE--PTVGEGLIVSMSLIGATIVTT 116
DB 31 AFACAMLASMTSILLGYDIGVMGAAIYIKRDLKVSDEQIEILLGINLYSLIGSC--- 87
QY 117 FSGPLSDSISGRPRMLLSILYFFSGLIMLSPNVYVLLARFVDGFGIGLAVTLVPLYI 176
DB 88 -AGRTSDWIGPRVITVIFAGTIFVIGALLMGFSFNYSFLMFGRFVAGIGIGYALMIAPVT 146
QY 177 SETAPSEIRGLNLTLPQFSGSGMFLSYCMWFGMS--LSPSPDWIRIMGLVLAIPSLFFGL 235
DB 147 AEVSPASSRGFLTSFFEVFINGGILIGIYSNAPSKLTILKVGWRMMLGVGALPSV-LLTV 205
QY 236 TIFVLPESPRMLVSGRMAEAKVLOKLRGKDDVSGELSLLEGLVGGDTSEIEYIIGP 295
DB 206 GVLAIMPESPRMLVNRGLGEARKVLNK----- 232
QY 296 ATEAADDLVDGDKQITLYGPEEQSWIARPSKPIMLGSVLASRHGSMVNSQVPLM 355
DB 233 -----TSDSKEE-----AQLRLA----- 245
QY 356 DPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFVTDQAKNEQWDEENLHRDDEYASD 415
DB 246 -----HIKQAAIGPESC-----NDD----- 260
```

QY 416 GAGGDYEDNLHSLRSQATGAEGKDI VHHGHRGSALEMRQETLLGEGDGVSTDI GGG 475
Db 261 -----VVQVNRQS-----NGEGV----- 273
QY 476 WQAWKWEKEGNGRKEGFKRVYLHQBGVPGSRGSI VSLPGGDVFESEFVHAAAL 535
Db 274 ----- 273
QY 536 VSQALFSKGLAEPMSDAAMVHPSEVAAGSRWKOLF---EPGVRALLVGVGIQILQQ 592
Db 274 -----WKELFLYPTPAIRHIVIAALGIIHFQQ 300
QY 593 FAGINGVLYTPQILEQAGVAVILSKFGLSSASASILISLTLMLPCIGFAMLLMDLS 652
Db 301 ASGVDAVLYSPRIFEXAGIINDTHKL-----LATVAVGFVKTVFILA-----ATFTLDRV 351
QY 653 GRFFLLGTIPILIASILVIVS-NLIDLG---TLAALLSTVS-VIVYFCFVWGFQPIIP 708
Db 352 GRPPLLSVGGWVLSLTLAISLTVIDHSERKLMWAGVSSIAMVLAYVATFSTIGAGFIT 411
QY 709 NILCAEIPTRVRLGICIAICAFTEWIGDII VTYSLPVMNAIGLAGVFSIYAVVCLISPV 768
Db 412 WYSEIEPLRLAQGAAGVAVNRITSVAVSMTFSLTRAITIGGAFFLYCGIATVGIW 471
QY 769 VFELKVPETKGMPLVIT-EFFVAGAKQAAKA 800
Db 472 PFYTVLPETRGKTLEDMEGSGFTFRSKSNASKA 504

RESULT 10
US-09-291-922-20
; Sequence 20, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; EARLIER FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Zea mays
US-09-291-922-20

Query Match 12.0%; Score 489.5; DB 3; Length 513;
Best Local Similarity 21.6%; Pred. No. 2.5e-39;
Matches 169; Conservative 106; Mismatches 200; Indels 307; Gaps 17;

QY 46 VTSDDI---LEDKMSG---AVLVAIVASIGNLQWDMNATIAAVALYIKKEPQLQNEPT 98
Db 1 MASDELAKAVEPRKKNVKTASICALIASMASVILGIDVNGGAAMYIKKDLNI-TDVQ 59
QY 99 VEGHIVMSLIGATIVTTFGSP-LSDSIGRRPMLILSSILYFFSGILMLSPNVYLLLA 157
Db 60 LEILIGILSLV--SLFGSFAGARTSDRIGRLTVVPAAVIFFVGSLLMLGFAVNYGMLNAG 117
QY 158 RFVDFGFGIAGVTLVPLVYISEIAPSEIRGLNLTLPQSGSGGFLSCVWFGMSLSP-SP 216
Db 118 RFVAGVGVGGMTAPYTAISPAASRGFLTTFPEVFINIGILLGLYLSNFAARPLHL 177
QY 217 DWRIMLGLVAIPSLFFFGTLTIFVLPESPRWLVSQGRMAEAKVKLQKRGDDVSGELSLL 276
Db 178 GWRVMAIGA VPS-GLLALLVFCMPESPRLVLKGLRADARAVLEX----- 222
QY 277 LEGLEVGDTSIEYIIGPATEAADLVDGDKBQITLYGPEBQSQWIARPSKGPIMLGS 336

Db 223 -----TSATPEAAERLA----- 235
QY 337 VLSLSRHSWMVNSVPLMDPIVTLFGSVHENMPOAGGSMRSTLFPFGSMFSVTDQAK 396
Db 236 ----- 235
QY 397 NEQWDEENLHRDEEVASDAGDYEDNLHSLRSQATGAEGKDI VHHGHRGSALEMR 456
Db 236 ----- 235
QY 457 QTLGEGDGVSTDI GGGWQLAWKWEKEGNGRKEGFKRVYLHQBGVPGSRGSI V 516
Db 236 -----DIKAA-----AGIPKGLDGDVVT 253
QY 517 LP-----GGDVFEGSEFVHAAALVSQALFSKGLAEPMSDAAMVHPSEVAAGSRWKOLF 572
Db 254 VPGKEGGGEL-----QVWKKL 270
QY 573 F---EPGVRALLVGVGIQILQOQFAGINGVLYTPQILEQAGVAVILSKFGLSSASASIL 629
Db 271 ILSPFAVERILLSAVGLHFFQOASGSDSVQYSARLFKSAGITDDNKLGLVTC 325
QY 630 ISSLTLLMLPCIGFAMLLMDLSGRFLLLTIPILIASILVIL-----VSNLIDLGLTA 684
Db 326 VGVTKTFIL-----VATFLDRAGERPLLLISTGGMIVSLICLGSGLTVAGHHPDTKVAM 381
QY 685 HALLSTVSIVVYFCFVWGFPIPNILCAEIPTRVRLGICIAICAFTEWIGDII VTYSLP 744
Db 382 AVALCIASTLSYIAFFSIGLGPITGVYTSIEIFFPQVRLGFAVGASNRVTSAVISMTFL 441
QY 745 VMLNAGLAGVFSIYAVVCLISFVFLKVPETKGMPLVITEFF-----AVGAKQAA 797
Db 442 SLSKAITTGGSPFLYSGIAAVAWVFFTCPLPETRGTLEEMKLFGLGMDTCMAEAEADA 501
QY 798 AK 799
Db 502 AK 503

RESULT 11
US-09-291-922-28
; Sequence 28, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 28
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-291-922-28

Query Match 12.0%; Score 489; DB 3; Length 529;
Best Local Similarity 20.8%; Pred. No. 3e-39;
Matches 165; Conservative 106; Mismatches 215; Indels 306; Gaps 14;

QY 24 LPSVVALPGLPASPSCSSQEPVTSDILEDKMSGAVLVAIVASIGNLQWDMNATIAA 83
Db 19 NAAALPEFGAVHPN-----KGNFYAFTCALCASMATVILGYDGVMSGA 65
QY 84 VLYIKKEPQLONE--PTVEGLIVMSLIGATIVTTFGSPISDSIGRRPMLILSILYFFS 141

Db 66 SLVTKRDLQITDVGLEIMMGLSVYALIGSFL-----GARTSDMVGRVTVTVFAAAIFNNG 121
Qy 142 GLIMLWPNVAVLLARFVDGFGIGLAVTLVPLVYISIAIPSEIRGLNLTLPQFSGSGMF 201
Db 122 SLIMGFAVNAVAVGFRVTVGIGVYALVAVPVYTPVPSASARGPLTSFTEVFINVIL 181
Qy 202 LSCVWFGMSLSP-SPDWRIMLVLAIPSLPFFGLTIFVLPSPRWLVSKRVAEAKVL 260
Db 182 LGVYSNFAFARLPVHLNWRVLMGLIGAVPSA-LLALAMVFGMPSPRWLVKRLADARVL 240
Qy 261 QKLRGKDDVSGELSLLLEGLVGGDTSEIEYIIGPATEAADLVDGDKQITLYGPEEG 320
Db 241 ----- 240
Qy 321 QSWIARPSKGPIMLGSLVSLASRHGSMVNSVPLMDPIVTLFGSVHNMFPAGGSMRSTL 380
Db 241 ----- 240
Qy 381 FPNFGSMFVTDQAKNEQWDEENLHRDDEYASDGAGDYEDNLHSPLLSRQATGAEGK 440
Db 241 -----AKTSDTPEEAVERLDQKAA----- 260
Qy 441 DIVHHGRGSALSRRQTLLEGEGDGVSTDIGGWLQAWKWKSEKENGKGGFKRVY 500
Db 261 ----- 260
Qy 501 LHQEGVPSRRGSIVSLPGGGEVPEGSEFVHAALVSQSALFSKGLAEPMSDAAMVHPS 560
Db 261 -----AGIPRELGDVVWMPK-----TKGQEKV----- 285
Qy 561 EVAAGSRWKDLF---EPGVRRALLVGVGIQILQQFAGINGVLYTTPQILEQAGVAVILS 617
Db 286 -----WKELIFSPTPAMRILALGIIHFHQATGSDSVLVSPRVFQSGAGITGDNH 337
Qy 618 KFLGSSASILISLTLMLPCIGFAMLMDSLRRLGTLGPIILIASLV-----IL 672
Db 338 LLG-----ATCAMGVKMTLFI-----VATFOLDVRGRPLLTSTAGLACLIGLGT 388
Qy 673 VVSNLIDGLTLAHLSTVSVVYFCCFVWGFGPIPNILCAEIPTRVGRGLCAICATF 732
Db 389 VGRHPDAKVPWAGLCIVSILAVSFSGIGLPLTSVTSEVFLVRLGFLALGTSN 448
Qy 733 WIGDIIVTSLFVMLNATGLAGVFSIYAVVCLISFVFLVFKPTEKGMPLVITEFF--- 789
Db 449 RVTSAVMSFSLSKAITIGSFYAGIAAGIWIPIPTIPETRGLPLEIGKLFQMT 508
Qy 790 ---AVCAKQAAK 799
Db 509 DTAVEAQDTATK 520

RESULT 12

US-09-291-922-30
; Sequence 30, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 30
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Beta vulgaris
US-09-291-922-30

Query Match 12.0%; Score 488; DB 3; Length 549;

Best Local Similarity 21.3%; Pred. No. 4e-39; Mismatches 194; Indels 298; Gaps 19;
Matches 164; Conservative 113;

Qy 41 SSOEPTVSDIILED-----KMSGAVLVAIVASIGNLLOGWDAITAAAVLYIKKEF 91
Db 10 SDPPPTASKVTADFDPLKPKPRNKFAFACATLASMTSVLLGDIYGVMSGAIYVKEDW 69
Qy 92 QLONEPTVEGLVSMISLIGATITVTFSGPLSDISIGRRPMLILSSILYPPSGILMLWSPNV 151
Db 70 HISD--TOIGVLVGLINLYCLFGSPAAGRTSDWIGRRYIVLAGAIFVFGALLMGFATNY 127
Qy 152 YVLLARFVDGFGIGLAVTLVPLVYISIAIPSEIRGLNLTLPQFSGSGMFLSYCMVFGMS 211
Db 128 AFLWGRFVTGIVGVYALVAVPVYTPVPSASARGPLTSFTEVFINVILGYSINLAFS 187
Qy 212 LSPSP-DWRIMLVLAIPSLPFFGLTIFVLPSPRWLVSKRMAEAKVLQKLRGKDDVS 270
Db 188 SLPTHLNWRVLMGLIGAVPSI-FLAIGVLAIMPSPRWLVKRLGDAKKVLR----- 238
Qy 271 GELSLLLEGLVGGDTSEIEYIIGPATEAADLVDGDKQITLYGPEEGQSWIARPSK 330
Db 239 -----ISD-----SPEAQ----- 247
Qy 331 PIMLGSLVSLASRHGSMVNSVPLMDPIVTLFGSVHNMFPAGGSMRSTLFPNFGSMFV 390
Db 248 ----- 247
Qy 391 TDHAKVQWDEENLHRDDEYASDGAGDYEDNLHSPLLSRQATGAEGKDIVHHGRGS 450
Db 248 -----L 248
Qy 451 ALSMRQTLLEGEGDGVSTDIGGWLQAWKWKSEKENGKGGFKRVYLHQEGVPGSR 510
Db 249 RUSEIKQT-----AGIPA-- 261
Qy 511 RGSIVLPGGDDVFEGSEFVHAALVSQSALFSKGLAEPMSDAAMVHSEVAKSRWK 570
Db 262 -----ECDE-----DIYKVKTKIKSGNAVWK 283
Qy 571 DLF---EPGVRRALLVGVGIQILQQFAGINGVLYTTPQILEQAGVAVILSKFLSSASAS 627
Db 284 ELFPNTPAVRAVIAIGIHFHQATGSDSVLVSPRVFQSGAGITNARKQL-----LAT 338
Qy 628 ILFSSITLMLPCIGFAMLMDSLRRLGTLGPIILIASLVILVVS-NLIDLG---T 682
Db 339 VAVGVKTLFI-----VATFOLDKYGRRPLLTSGVGMIIAILTLAMSLTVIDHSHKIT 394
Qy 683 LAHALLSTVSVVYFCCFVWGFGPIPNILCAEIPTRV--GLCIAICAFTEWIGDIIVT 740
Db 395 WATALCITM-VCAVASFSGIGLPIVWYSSEVFLRLRAQGTSMGV-AVNRVVSVISI 452
Qy 741 YSLPVMNATGLAGVFSIYAVVCLISFVFLVFKPTEKGMPLVITEFF 789
Db 453 FFLP-LSHKITTGGAFFLGGIAI IAWFFLFLPETRGRITLENMHELF 500

RESULT 13

US-09-291-922-22
; Sequence 22, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998

; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 22
 ; LENGTH: 510
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; NAME/KEY: UNSURE
 ; LOCATION: (102)
 US-09-291-922-22

Query Match 11.9%; Score 485; DB 3; Length 510;
 Best Local Similarity 22.8%; Pred. No. 7e-39;
 Matches 173; Conservative 98; Mismatches 187; Indels 300; Gaps 19;

QY 59 AVLVAIVASIGNLQGDWNTAAAVLYIKKEFOLQNEPTVE--GLIVSMLIGATV 115
 DB 21 AFACILASMTSILGYDIGVMSGASLYIKKDFNI--SDGRVEVLGMLNLYSLIGSPA-- 77
 QY 116 TFGSPLSDSIGRRPMLILSSILYFFSGLIMLWSPNVYVLLARVDFGIGLAVTLVPLY 175
 DB 78 --AGRTSDWIGRYTIVFAAVIFPAGXFLMGFAVNYAMLFGFRVAGIGVGYALMIAPVY 135
 QY 176 ISETAPSEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSP--SPDWRIMLGVLAPSLFFPG 234
 DB 136 TABVSPASARGFLTSFPEVFINFGILLYGVSNAFSLPLNGLWRIMLGIGAAPS--LLA 194
 QY 235 LTIFVLPSPRLWYSKGRMAEAKVLOKLRGKDDVSGELSLLLEGLVGGDTSTEEYIIG 294
 DB 195 LMLVGLMPSPRLWYKGLADAKVLEKT-----SDT----- 226
 QY 295 PATEAADLVDGDKQETITLYGPEGQSWIARPSKGPIMLGVSLSLASRSGSMWNQSVPL 354
 DB 227 -AEEAERL----- 234
 QY 355 MDPIVTLFGSVHNNPQAGSMRSTLPNFGSMFVTDQAKNQDEENLRHDEBYAS 414
 DB 235 ----- 234
 QY 415 DGAGDYEDNLHSLPSSRQATGAEKDIIVHHRGSALSMMRQTLGEGDGVSTDIGG 474
 DB 235 -----ADIK 239
 QY 475 GWQLAWKSEKENGKEGFKVYLHQEVGVRSGSVSLP--CGGDVFGSEFVHA 532
 DB 240 A-----AGIPEELGDVVTVPKRSGN----- 261
 QY 533 AALVQSALFSKGLAEPRMSDAMVHPSEVAAGSRWKDLF---EPGVRALLVGVGIQI 589
 DB 262 -----EKRV-----WKELISPTPAVRILLSGIGIHF 289
 QY 590 LQFAGINGVLYYTPQILEQAVAVILSKGLSSASASILSSITLMLPCPGFAMLM 649
 DB 290 FOHALGHVSVFPLVFKSPGLT--NDKHFLGTWTFPGVTKRFLILL-----ATFFI 340
 QY 650 DLGRRFPALLGTIPILIASLV-----ILVVSNDLIDGLTAHALISTVSVIVFCFVMGF 704
 DB 341 DGVGRPLLGSTGIIILSLIGLAGLTVVQHPDAKI PWAIGLSIASTLAYVAFSIGL 400
 QY 705 GPIINILCAEIPPRVR--GLCINCAFTWIGDIIVTYSPLPMLNAGLAGVFSYAVV 762
 DB 401 GPIITWYSSEIFPLQVRALGCSLGVANRVTSVVISMTFL--SLSKAITGGFFLYSGI 458
 QY 763 CLISFVFLKPRRTKMPLEVIETFEFAVGAQAQAAKA 800
 DB 459 AALAWVFYVTLPTRTGLEEMSKLP--GDTAAASES 494

RESULT 14
 US-09-291-922-26
 ; Sequence 26, Application US/09291922
 ; Patent No. 6383776
 ; GENERAL INFORMATION:

; APPLICANT: Allen, Steve
 ; APPLICANT: Hitz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB-1163
 ; CURRENT APPLICATION NUMBER: US/09/291,922
 ; CURRENT FILING DATE: 1999-04-14
 ; EARLIER APPLICATION NUMBER: 60/083,044
 ; EARLIER FILING DATE: April 24, 1998
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 26
 ; LENGTH: 539
 ; TYPE: PRT
 ; ORGANISM: Triticum aestivum
 US-09-291-922-26

Query Match 11.4%; Score 463.5; DB 3; Length 539;
 Best Local Similarity 21.7%; Pred. No. 1.1e-36;
 Matches 166; Conservative 108; Mismatches 200; Indels 291; Gaps 17;

QY 52 LEDKMSG----AVLVAIVASIGNLQGDWNTAAAVLYIKKEFOLQNEPTVE---GLIV 104
 DB 35 VEPKKGWVRFAFACILASMTSILGYDIGVMSGASLYIKDLKI--NDTQLEVLGMLN 93
 QY 105 SMSLICATITVTFSGPLSDSISGRPRMLILSSILYFFSGLIMLWSPNVYVLLARVDFG 164
 DB 94 VYSLIGSPA---AGRTSDWIGRYTIVFAAVIFPAGALINGFVNYAMLFGFRVAGIG 149
 QY 165 IGLAVTLPLYISETAPSEIRGLNLTLPQFSGSGMFLSYCMVFGMS--LSPSDWRIMLG 223
 DB 150 VGVALIAPVNTGEVSPASARGVLTSPFPEVFINFGILLYGVSNAFARLSLRLGRLM 209
 QY 224 VLAIPSLPFFGLTIFVLPSPRLWYSKGRMAEAKVLOKLRGKDDVSGELSLLLEGLV 283
 DB 210 IGAVFSV--LLAPVLMGPSPRLWYKGLADAKVVLAKT----- 248
 QY 284 GTSIEEYIIGPATEAADLVDGDKQETITLYGPEGQSWIARPSKGPIMLGVSLSLASR 343
 DB 249 SDT-----PEAAERIA----- 260
 QY 344 HSMWNQSVPLMDPIVTLFGSVHNNPQAGSMRSTLPNFGSMFVTDQAKNEQWDEE 403
 DB 261 ----- 260
 QY 404 NLHRDEYASDGAGDYEDNLHSLPSSRQATGAEKDIIVHHRGSALSMMRQTLGEG 463
 DB 261 -----DIKTAA----- 266
 QY 464 DGVVSTDIQGGWQLAWKSEKENGKEGFKVYLHQEVGVRSGSVSLPGGGDV 523
 DB 267 -----GIPLDGDVVTVPK----- 281
 QY 524 FEGSEFVHAALVQSALFSKGLAEPRMSDAMVHPSEVAAGSRWKDLFEPGVRALLV 583
 DB 282 -----NKGSEKRVKDLILSPITIA-----NRHILIA 309
 QY 584 GVGIILOQFAGINGVLYYTPQILEQAVAVILSKFGLSS--ASASILSSITLMLPCI 642
 DB 310 GIGIHFOQSGIDAVVLYSPLVFKSAGIT-----GDSRLRGTTVAGTNTVFIL--- 360
 QY 643 GFAMLMIDLGRFPALLGTIPILIASLV-----ILVVSNDLIDGLTAHALISTVSVIV 697
 DB 361 -VATELLDRIRRRPLVLTSTGMLVSLVGLATGLTVISRHPDEKITWAILVICIMAYV 419
 QY 698 CCFWNGPFIPIINILCAEIPPRVRGLCTAICAFWTFWIGDIIVTYSPLPMLNAGLAGVFS 757
 DB 420 AFFSIGLGPITWYSSEIFPLHVRALGCSLGVANRVLTSGVISMTFTLSLKAITGG 479
 QY 758 IVAVVCLISFVFLKPRRTKMPLEVIETFEFAVGA--KQAAAKA 800
 DB 480 LFAGIASFAWVFYVTLPTRTGLEEMSKLP--GDTAAASES 524

OY 746 MNAIGLAGVRSIYAVVCLISFVFLXVPETKGMPLVI 785
Db 420 MLNSIGSANTFWYGLNVLFIITLWLPETKNVSLSHI 459

Search completed: October 13, 2004, 11:40:10
Job time : 30.8337 secs

RESULT 15
US-09-489-039A-11933
; Sequence 11933 Application US/09489039A
; Patent No. 8610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11933
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11933

Query Match 11.0%; Score 447; DB 4; Length 476;
Best Local Similarity 20.8%; Pred. No. 3.7e-35;
Matches 158; Conservative 98; Mismatches 186; Indels 318; Gaps 18;

OY 42 SOEPVTS-DDILEP-KMSG-----AVLVAIVASIGNLOGWONATIAAAVLYIKKEFOL 93
Db 2 STESITOLEGVPNDKKQGRSNKMTTFVCFLAALAGLLFGLDGVAGALPTIANEFQI 61

OY 94 QNEPTVEGLIVMSLIGATIVTTPSGPLSDSIGRRPMLILSSILYPPSGLIMLSPNVYV 153
Db 62 SAH--TQEWVSSMMFGAAGVAGSGWLSFKLGRKKSIMIGAILFVAGSLFSAAPNVEI 119

OY 154 LLARFVDGFGIGLAVTLPIYIETAPSEIRGLNLTLPQPSGCGMFLSCWVFGMSLS 213
Db 120 LLVSRVLLGLAVGASVATPYLSEIAPKEIRGSMISMYQIMITIGILGAY--LSDTAFS 177

OY 214 PSPWRIMLGVLTIPSPFFGLTIFYPSPRWLVSKRMAEAKVKLQKRGDDVSGEL 273
Db 178 YSGARWMLGVIIIPAVULL-IGVIFLPSRWFAAKRFVDAERVLRLR----- 227

OY 274 SLLLEGEVGGDTSEIEYIIGPATEADDLVTDGDKQKITLYGPEGQSWIARPSKGPIM 333
Db 228 -----DTSAE----- 232

OY 334 LGSVLSLASRHGSMVNSQSVPLMDPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFSVTDQ 393
Db 233 ----- 232

OY 394 HAKNEQWDEENLHRDEEYASDGAGDYEDNLHSPILSRQATGAEGKDIVHHGRGSALS 453
Db 233 -AKREL-----DEIRESLKVKQS----- 249

OY 454 MRQTLLEGDDGVSSSTDIGGWQLAWKWKSEKEGNGRKEGFKRVYLHQEGVFGSRGS 513
Db 250 -----GWSL-----FK----- 255

OY 514 IVSLPGGDVPEGSFVHAALVQSALFSKGLAEPMSDAAMVHPSEVAAGSRWKDLF 573
Db 256 -----DNSNF----- 260

OY 574 EPGVERALLVVGIGIQQOQFAGINGVLYYTPQLEQAGVAVILSKFGLSSASAILISL 633
Db 261 ----RRAVFLGILLQVMQOFTGMNVIMYAPKIFELAGYANTTEQM-----GTIVGLT 311

OY 634 TTLMLPCIGFAMLLDLSGRFLLGTPTILIASLIVLVVSNLIDIGTLAH----- 685
Db 312 NVLATFIAIG-----LVDWRGRK-----PTILGFIWAAAGMV-LGTMHGHHSSTA 359

OY 686 ALLSTVSVIVYPCCFVMGFGPIPNILCAEIPFTRVRLGICIAICAFTFWIGDIIVTYSLPV 745
Db 360 QYIAVLMMLMFIVGAMSAGPLIWLCSLQKGRDGRGIICTCATNWIANNIWGATELT 419

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 11:38:38 ; Search time 302.735 Seconds
(without alignments)
852.715 Million cell updates/sec

US-10-051-909-32

Title: 4075
Perfect score: 4075
Sequence: 1 IRSGSMIAVQTPTDRLDR.....PLEVITEFFAAGKQAQAAKA 800

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1356558 seqs, 322682953 residues

Total number of hits satisfying chosen parameters: 1356558

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4075	100.0	800	13	US-10-051-909-32
2	3774	92.6	745	14	Sequence 32, Appl
3	3517	86.3	747	13	Sequence 718, App
4	3517	86.3	747	13	Sequence 2, Appl
5	2869	70.4	564	15	Sequence 2, Appl
6	2834	69.5	740	13	Sequence 56898, A
7	2834	69.5	740	16	Sequence 37, Appl
8	2674	65.6	737	13	Sequence 195955
9	2674	65.6	737	13	Sequence 8, Appl
10	2584	63.4	729	13	Sequence 8, Appl
11	2584	63.4	729	13	Sequence 2, Appl
12	2581.5	63.3	742	15	Sequence 153166
13	2543	62.4	728	15	Sequence 47258, A
14	2496.5	61.3	742	15	Sequence 251277

15	2442.5	59.9	661	15	US-10-425-114-37283	Sequence 37283, A
16	2305	56.6	735	16	US-10-437-963-149907	Sequence 149907
17	1963	48.2	652	16	US-10-437-963-195442	Sequence 195442
18	1763	43.3	714	16	US-10-437-963-173392	Sequence 173392
19	1670.5	41.0	486	13	US-10-051-902-10	Sequence 10, Appl
20	1670.5	41.0	486	13	US-10-051-909-10	Sequence 10, Appl
21	1557	38.2	648	15	US-10-424-599-220809	Sequence 220809
22	1424	34.9	345	13	US-10-051-902-14	Sequence 14, Appl
23	1424	34.9	345	13	US-10-051-909-14	Sequence 14, Appl
24	1293	31.7	272	16	US-10-767-701-44159	Sequence 44159, A
25	1252.5	30.7	338	15	US-10-424-599-184842	Sequence 184842
26	944	23.2	248	15	US-10-424-599-243978	Sequence 243978
27	943	23.1	228	13	US-10-051-902-16	Sequence 16, Appl
28	943	23.1	228	13	US-10-051-909-16	Sequence 16, Appl
29	813	20.0	161	15	US-10-425-114-69883	Sequence 69883, A
30	793	19.5	158	15	US-10-425-114-47887	Sequence 47887, A
31	663	16.3	145	16	US-10-767-701-50386	Sequence 50386
32	626	15.4	218	16	US-10-437-963-162571	Sequence 162571
33	626	15.4	218	16	US-10-767-701-44178	Sequence 44178, A
34	587.5	14.4	457	14	US-10-369-493-23324	Sequence 23324, A
35	562	13.8	131	13	US-10-051-902-6	Sequence 6, Appl
36	562	13.8	131	13	US-10-051-909-6	Sequence 6, Appl
37	554.5	13.6	487	9	US-09-795-693-27	Sequence 27, Appl
38	554.5	13.6	487	13	US-10-095-139-14	Sequence 14, Appl
39	554.5	13.6	487	14	US-10-156-239-27	Sequence 27, Appl
40	554.5	13.6	487	14	US-10-199-485-27	Sequence 27, Appl
41	554.5	13.6	487	14	US-10-391-399-45	Sequence 45, Appl
42	554	13.6	143	15	US-10-425-114-55090	Sequence 55090, A
43	548	13.5	501	15	US-10-424-599-239553	Sequence 239553
44	548	13.4	488	10	US-09-794-822-11	Sequence 11, Appl
45	548	13.4	488	13	US-10-094-059-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-10-051-909-32
; Sequence 32, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Heientjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: B31163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 32
; LENGTH: 800
; TYPE: PRT
; ORGANISM: Zea mays
US-10-051-909-32

Query Match	100.0%	Score 4075;	DB 13;	Length 800;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches	800;	Conservative 0;	Mismatches 0;	Gaps 0;
Qy	1	IRSGSMIAVQTPTDRLDRRLPSVLLPGLPPASCSQEPVTSDDILEDKMSGAV	60	
Db	1	IRSGSMIAVQTPTDRLDRRLPSVLLPGLPPASCSQEPVTSDDILEDKMSGAV	60	
Qy	61	LVAIVASIGNLQGNWNTAAAVLYKKEFQLEQNEPTVEGLIVSMGLIGATVITTFSGP	120	
Db	61	LVAIVASIGNLQGNWNTAAAVLYKKEFQLEQNEPTVEGLIVSMGLIGATVITTFSGP	120	
Qy	121	LSDSIGRRPMLLSSILYFFSGIMLWSPNVYLLARFVDGFGIGLAVTLVPLYISEIA	180	

Db 121 LSDSLGRPMILLSILIFFSGLIMLWSPNYVLLARFVDFGIGLAVTLVPLVISEIA 180
 QY 181 PSEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPSPDWRLMGLVLAIPSLFFGLTIFYL 240
 Db 181 PSEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPSPDWRLMGLVLAIPSLFFGLTIFYL 240
 QY 241 PSEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPSPDWRLMGLVLAIPSLFFGLTIFYL 240
 Db 241 PSEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPSPDWRLMGLVLAIPSLFFGLTIFYL 240
 QY 301 DDVLTGDKQEQITLYGPEGQSWIARPSKGPIMLGSVLSLASRHGSMVNSQVPLMDPIVT 360
 Db 301 DDVLTGDKQEQITLYGPEGQSWIARPSKGPIMLGSVLSLASRHGSMVNSQVPLMDPIVT 360
 QY 361 LFGSVHNMPOAGSGMRSTLFPNFGSMFVTDQHAKNEQWDEENLHRDDEEYASDGAGD 420
 Db 361 LFGSVHNMPOAGSGMRSTLFPNFGSMFVTDQHAKNEQWDEENLHRDDEEYASDGAGD 420
 QY 421 YEDNLHSPLLSRQATGAEKDIIVHGHGSGALSMMRQTLGEGGDSVSTTDIGGQWOLAW 480
 Db 421 YEDNLHSPLLSRQATGAEKDIIVHGHGSGALSMMRQTLGEGGDSVSTTDIGGQWOLAW 480
 QY 481 KWEKEGNGRKEGKFKVYVHQBGVPGSRGSIIVSLPGGSDVPEGSEFVHAAALVSQSA 540
 Db 481 KWEKEGNGRKEGKFKVYVHQBGVPGSRGSIIVSLPGGSDVPEGSEFVHAAALVSQSA 540
 QY 541 LFSKGLARPRMSDAWHPSEVAAGSRWKDLFPBGVRRALLVGVGQIQQFAGINGVL 600
 Db 541 LFSKGLARPRMSDAWHPSEVAAGSRWKDLFPBGVRRALLVGVGQIQQFAGINGVL 600
 QY 601 YTPQILBQAGVAVILSKFGLSSASASILSSLTLLMLPCIGFAMLLMDLSGRFRLLG 660
 Db 601 YTPQILBQAGVAVILSKFGLSSASASILSSLTLLMLPCIGFAMLLMDLSGRFRLLG 660
 QY 661 TIPILIASLVILVNSLIDLTGTAHALLSTVSVIVYCCFVWVGPIPNILCAEIPFTRV 720
 Db 661 TIPILIASLVILVNSLIDLTGTAHALLSTVSVIVYCCFVWVGPIPNILCAEIPFTRV 720
 QY 721 RGLCIAICAFTEWIGDIIVTYSLPVMLNAILAGVFSIYAVVCLISFVFLKVPETKGM 780
 Db 721 RGLCIAICAFTEWIGDIIVTYSLPVMLNAILAGVFSIYAVVCLISFVFLKVPETKGM 780
 QY 781 PLEVITEFFAVGAKAAKA 800
 Db 781 PLEVITEFFAVGAKAAKA 800

RESULT 2
 US-10-310-154-718

; Sequence 18 Application US/10310154
 ; Publication No. US20030233670A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Edgerton, Michael D
 ; APPLICANT: Chomet, Paul S.
 ; APPLICANT: Adams, Thomas H
 ; APPLICANT: Ruff, Thomas G.
 ; APPLICANT: Ahrens, Jeffrey E.
 ; APPLICANT: Ball, James A.
 ; APPLICANT: Banu, G.
 ; APPLICANT: Bell, Erin
 ; APPLICANT: Bodupalli, Raghava
 ; APPLICANT: Deikman, Jill
 ; APPLICANT: Deng, Molian
 ; APPLICANT: Dong, Jinhao
 ; APPLICANT: Duff, Stephen M.
 ; APPLICANT: Galligan, Meghan M.
 ; APPLICANT: Hinchey, Brenda S.
 ; APPLICANT: Huang, Shishieh
 ; APPLICANT: Johnson, G. Richard
 ; APPLICANT: Jung, Vincent
 ; APPLICANT: Kretzmer, Keith A

; APPLICANT: Laccetti, Lucille B.
 ; APPLICANT: Lai, Chao-Qiang
 ; APPLICANT: Lee, Gary
 ; APPLICANT: Lin, Jie-Yi
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Lu, Bin
 ; APPLICANT: Luethy, Michael M.
 ; APPLICANT: Lund, Adrian
 ; APPLICANT: Madison, Linda L.
 ; APPLICANT: Malloy, Kathleen A.
 ; APPLICANT: McKiel, Christine L.
 ; APPLICANT: Miller, Philip W.
 ; APPLICANT: Padmavathi, Manchikanti
 ; APPLICANT: Parnell, Laurence D.
 ; APPLICANT: Start, William G.
 ; APPLICANT: Tennessee, Dan
 ; APPLICANT: Vidya, K.R.
 ; APPLICANT: Wang, Haiyun
 ; APPLICANT: Xin, Zhanguo
 ; APPLICANT: Xu, Nanfei
 ; APPLICANT: Yang, Chunzhi
 ; APPLICANT: Zeng, Xiaoping
 ; APPLICANT: Zhang, Qiang
 ; APPLICANT: Zhao, Yajuan
 ; APPLICANT: Zhou, Li
 ; TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
 ; FILE REFERENCE: 38-15(52796)B
 ; CURRENT APPLICATION NUMBER: US/10/310,154
 ; CURRENT FILING DATE: 2002-12-04
 ; PRIOR APPLICATION NUMBER: 60/337,359
 ; PRIOR FILING DATE: 2001-12-04
 ; NUMBER OF SEQ ID NOS: 736
 ; SEQ ID NO 718
 ; LENGTH: 745
 ; TYPE: PPT
 ; ORGANISM: Zea mays
 ; US-10-310-154-718
 Query Match 92.6%; Score 3774; DB 14; Length 745;
 Best Local Similarity 99.6%; Pred. No. 1e-315;
 Matches 742; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 56 MSGAVLVAIVASIGNLQGNATIAAAVLYIKKEFQLONEPTVEGLIVNSLIGATIVT 115
 Db 1 MSGAVLVAIVASIGNLQGNATIAAAVLYIKKEFQLONEPTVEGLIVNSLIGATIVT 60
 QY 116 TFSGLSDSIGRRPMLIILSSILYFFSGLIMLWSPNYVLLARFVDFGIGLAVTLVPLV 175
 Db 61 TFSGLSDSIGRRPMLIILSSILYFFSGLIMLWSPNYVLLARFVDFGIGLAVTLVPLV 120
 QY 176 ISEIAPSEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPSPDWRLMGLVLAIPSLFFGL 235
 Db 121 ISEIAPSEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPSPDWRLMGLVLAIPSLFFGL 180
 QY 236 TIFYPESPRLVSKGRMAEAKVLOKRGKDDVSGELSLLEGLVGGDTSEIYEIIGP 295
 Db 181 TIFYPESPRLVSKGRMAEAKVLOKRGKDDVSGELSLLEGLVGGDTSEIYEIIGP 240
 QY 296 ATEAADDLVTDGDKQEQITLYGPEGQSWIARPSKGPIMLGSVLSLASRHGSMVNSQVPLM 355
 Db 241 ATEAADDLVTDGDKQEQITLYGPEGQSWIARPSKGPIMLGSVLSLASRHGSMVNSQVPLM 300
 QY 356 DPVTLFGSVHNMPOAGSGMRSTLFPNFGSMFVTDQHAKNEQWDEENLHRDDEEYASD 415
 Db 301 DPVTLFGSVHNMPOAGSGMRSTLFPNFGSMFVTDQHAKNEQWDEENLHRDDEEYASD 360
 QY 416 GAGDYEDNLHSPLLSRQATGAEKDIIVHGHGSGALSMMRQTLGEGGDSVSTTDIGG 475
 Db 361 GAGDYEDNLHSPLLSRQATGAEKDIIVHGHGSGALSMMRQTLGEGGDSVSTTDIGG 420
 QY 476 WOLAKWSEKEGNGRKEGKFKVYVHQBGVPGSRGSIIVSLPGGSDVPEGSEFVHAAAL 535
 Db 421 WOLAKWSEKEGNGRKEGKFKVYVHQBGVPGSRGSIIVSLPGGSDVPEGSEFVHAAAL 480

```
QY 536 VSQALFSKGLAEPMSDAAMVHSEVAAGSRWKDLFEPGVRRALLVGVGIQLQOFAG 595
DB 481 VSQALFSKGLAEPMSDAAMVHSEVAAGSRWKDLFEPGVRRALLVGVGIQLQOFAG 540
QY 596 INGVLYYTPQILEQAGVAVILSKFGLSSASASIISSITLLMLPCIGFAMLLMDLSGRR 655
DB 541 INGVLYYTPQILEQAGVAVILSKFGLSSASASIISSITLLMLPCIGFAMLLMDLSGRR 600
QY 656 FLLGTTPIILIASVILVWSNLDLGTALHALLSTVSVIVYFCFVNGFGPIPNILCAEI 715
DB 601 FLLGTTPIILIASVILVWSNLDLGTALHALLSTVSVIVYFCFVNGFGPIPNILCAEI 660
QY 716 FPTRVGLCIAICAFTEWIGDIIIVTYSLPVNLNAIGLAGVFSIYAVVCLISFVFLKVP 775
DB 661 FPTRVGLCIAICAFTEWIGDIIIVTYSLPVNLNAIGLAGVFSIYAVVCLISFVFLKVP 720
QY 776 ETKCMPLVITEFFAVGAKQAAKA 800
DB 721 ETKCMPLVITEFFAVGAKQAAKA 745

RESULT 3
US-10-051-902-2
; Sequence 2, Application US/10051902
; Publication No. US20020178469A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 747
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (129)
; NAME/KEY: UNSURE
; LOCATION: (133)..(134)
; NAME/KEY: UNSURE
; LOCATION: (144)
; NAME/KEY: UNSURE
; LOCATION: (178)
; NAME/KEY: UNSURE
; LOCATION: (207)
; NAME/KEY: UNSURE
; LOCATION: (218)
; NAME/KEY: UNSURE
; LOCATION: (220)
; NAME/KEY: UNSURE
; LOCATION: (236)
US-10-051-902-2

Query Match 86.3%; Score 3517; DB 13; Length 747;
Best Local Similarity 91.6%; Pred. No. 1.4e-293;
Matches 684; Conservative 27; Mismatches 34; Indels 2; Gaps 2;

QY 56 MSGAVLVAIVASIGNLLQGDWNTAIAAVLYIKKEFQLQNEPTVEGLIVMSLIGATVIT 115
DB 1 MGGAVVVAIAASIGNLLQGDWNTAIAAGAVLYIKKEFNQSEPLIEGLIVAMFLIGATVIT 60
QY 116 TSPGSLSDSISGRPMMLIISSILYFSGGLIMLWSPNVYLLILARFVDFGFGIATVLPY 175
DB 61 TSPGPRADCVGRPMPLVASAVLYFVSGGLVWMLWAPIVYILLARLIDGFGIATVLPY 120
```

```
QY 176 ISEIAPSEIR-GUANTLPOFSG-SGOMFLSYCMVFGMSLSPSPDWIRIMLVLAISLPFF 233
DB 121 ISETAFHRSXWXXNTLPQFIGVXGMFLSYCMVFGMSLMPKPDWRLMLGVLSISLXYF 180
QY 234 GLTIFYPSPRWLVSKGRMAEAKVYLQKLRKODVSGELSLLLLLEGLVGGTSTEEYII 293
DB 181 GLTVFYPSPRWLVSKGRMAEAKRVXQRLRGREDVSEXALLVEGLGVGKDTRIXEYII 240
QY 294 GPATEAADLVTDGKEQITLYGPBEGOSWIARPSKGPIMLGSLVLSASRHGSMVNSQVP 353
DB 241 GPATEAADLVTDGKEQITLYGPBEGOSWIARPSKGPIMLGSLVLSASRHGSMVNSQVP 300
QY 354 LMDPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFVTDQHAKNPCWBEENLHRDDEYA 413
DB 301 LMDPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFVTDQHAKNPCWBEENLHRDDEYA 360
QY 414 SDGAGDYEDNLHSPILLSQATGAGKQIVHKGHSGSALSMSRQTLLEGGDGVSTDIG 473
DB 361 SDGAGDYEDNLHSPILLSQATGAGKQIVHKGHSGSALSMSRQTLLEGGDGVSTDIG 420
QY 474 GGQQLAWKWKSEKENGKRGEGFKRVYLHQEGVPGSRSGSIIVSLPGGSDVFESEFVHAA 533
DB 421 GGQQLAWKWKSEKENGKRGEGFKRVYLHQEGVPGSRSGSIIVSLPGGSDVFESEFVHAA 480
QY 534 ALVQSALFSKGLAEPMSDAAMVHSEVAAGSRWKDLFEPGVRRALLVGVGIQLQOF 593
DB 481 ALVQSALFSKGLAEPMSDAAMVHSEVAAGSRWKDLFEPGVRRALLVGVGIQLQOF 540
QY 594 AGINGVLYYTPQILEQAGVAVILSKFGLSSASASIISSITLLMLPCIGFAMLLMDLSG 653
DB 541 AGINGVLYYTPQILEQAGVAVILSKFGLSSASASIISSITLLMLPCIGFAMLLMDLSG 600
QY 654 RRFLLLGTTPIILIASVILVWSNLDLGTALHALLSTVSVIVYFCFVNGFGPIPNILCA 713
DB 601 RRFLLLGTTPIILIASVILVWSNLDLGTALHALLSTVSVIVYFCFVNGFGPIPNILCA 660
QY 714 EIPPTRVGLCIAICAFTEWIGDIIIVTYSLPVNLNAIGLAGVFSIYAVVCLISFVFLK 773
DB 661 EIPPTRVGLCIAICAFTEWIGDIIIVTYSLPVNLNAIGLAGVFSIYAVVCLISFVFLK 720
QY 774 VPETKCMPLVITEFFAVGAKQAAKA 800
DB 721 VPETKCMPLVITEFFAVGAKQAAKA 747

RESULT 4
US-10-051-909-2
; Sequence 2, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 747
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (129)
; NAME/KEY: UNSURE
; LOCATION: (133)..(134)
```

```

; NAME/KEY: UNSURE
; LOCATION: (144)
; NAME/KEY: UNSURE
; LOCATION: (178)
; NAME/KEY: UNSURE
; LOCATION: (207)
; NAME/KEY: UNSURE
; LOCATION: (218)
; NAME/KEY: UNSURE
; LOCATION: (220)
; NAME/KEY: UNSURE
; LOCATION: (236)
US-10-051-909-2

Query Match      86.3%; Score 3517; DB 13; Length 747;
Best Local Similarity 91.6%; Pred. No. 1.4e-253;
Matches 684; Conservative 27; Mismatches 34; Indels 2; Gaps 2;

QY 56 MSGAVLVAIVASIGNLLQGDWNTAATAAVLYIKKEFOLONEPTVEGLIVSMISLIGATIVT 115
DB 1 MGGAVWVAIAASIGNLLQGDWNTAIAAGVLYIKKEFNLOSEPLIEGLIVAMFLIGATIVT 60

QY 116 TFSGPLSDISGRRPMLILSSILYFPGSLIMLSPNVYVLLIAREVDGFGIGLAVTLVPLY 175
DB 61 TSPGPRADCVGRPRPMLVASAVLYFVSGLVMLWAPIVILLIARLIDFGIGLAVTLVPLY 120

QY 176 ISETAPSEIR-GLANTLPOPSG-SGGMFLSCYCVFGMSLSPSPDWRMLGVLAIPSLFF 233
DB 121 ISETAPHXSGXNTLPOQFVGVGGMFLSCYCVFGMSLMPKPDWRMLGVLSIPSLXYF 180

QY 234 GLTIFLPSPRWLVSXGMAEAKVLLQKRGKDDVSGELSLLEGLVGGDTSEIEYII 293
DB 181 GLTVFYLPSPRWLVSXGMAEAKVXQRLRGREDVSKXALLVEGLGVGKDTREXEYII 240

QY 294 GPATEAADLVTGDKKEQITLYGPEEGSWIARPSKGPIMLGSVLSLASRHGSMVQSV 353
DB 241 GPATEAADLVTGDKKEQITLYGPEEGSWIARPSKGPIMLGSVLSLASRHGSMVQSV 300

QY 354 LMDPLVTLFGSVHENMPQAGSMRSTLFPNFGSMFVTDQAKNEQWDEENLHRDDEEYA 413
DB 301 LMDPLVTLFGSVHENMPQAGSMRSTLFPNFGSMFVTDQAKNEQWDEENLHRDDEEYA 360

QY 414 SDGAGGYEDNLHSPILSRQATGAEGKDI VHHGHRGSGALSMRQTLLEGEGDGVSTDIG 473
DB 361 SDGAGGYEDNLHSPILSRQATGAEGKDI VHHGHRGSGALSMRQTLLEGEGDGVSTDIG 420

QY 474 GGWQLAWKWSKEGNGRKEGGFKRVYLHQBGVPGSRGSI VSLPGGDDVFESEFVHAA 533
DB 421 GGWQLAWKWSKEGNGRKEGGFKRVYLHQBGVPGSRGSI VSLPGGDDVFESEFVHAA 480

QY 534 ALVSCSALFSKGLAEPMSDAAMVHPSEVAAGSRWKDLPEPGVRRALLVGVGIQLQQF 593
DB 481 ALVSCSALFSKGLAEPMSDAAMVHPSEVAAGSRWKDLPEPGVRRALLVGVGIQLQQF 540

QY 594 AGINGVLYTTPQILEQAGVAVILSKFGLSSASILSSLTTLMLPCIGFAMLLDLSG 653
DB 541 AGINGVLYTTPQILEQAGVAVILSKFGLSSASILSSLTTLMLPCIGFAMLLDLSG 600

QY 654 RRFLLGTIPILIASLVILVNSNLDLGTLAHALLSTVSVIVFCCFVMGFGPIPNILCA 713
DB 601 RRFLLGTIPILIASLVILVNSNLDLGTLAHALLSTVSVIVFCCFVMGFGPIPNILCA 660

QY 714 EIPFTRVGLCIAICATFTWIGDIIIVTYSLPVMLNAICLAGVFSIYAVVCLISFVFLK 773
DB 661 EIPFTRVGLCIAICATFTWIGDIIIVTYSLPVMLNAICLAGVFSIYAVVCLISFVFLK 720

QY 774 VPETKGMPLVITEFFFAVGAQAAAKA 800
DB 721 VPETKGMPLVITEFFFAVGAQAAAKA 747

```

RESULT 5
US-10-425-114-56898

```

; Sequence 56898, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 56898
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLM017045C09_FLI.pep
US-10-425-114-56898

Query Match      70.4%; Score 2869; DB 15; Length 564;
Best Local Similarity 99.5%; Pred. No. 6.2e-238;
Matches 561; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 237 IFYLPESPRWLVSXGMAEAKVLLQKRGKDDVSGELSLLEGLVGGDTSEIEYII 296
DB 1 IFYLPESPRWLVSXGMAEAKVLLQKRGKDDVSGELSLLEGLVGGDTSEIEYII 60

QY 297 TEAADLVTGDKKEQITLYGPEEGSWIARPSKGPIMLGSVLSLASRHGSMVQSVPLMD 356
DB 61 TEAADLVTGDKKEQITLYGPEEGSWIARPSKGPIMLGSVLSLASRHGSMVQSVPLMD 120

QY 357 PIVTLFGSVHENMPQAGSMRSTLFPNFGSMFVTDQAKNEQWDEENLHRDDEEYASDG 416
DB 121 PIVTLFGSVHENMPQAGSMRSTLFPNFGSMFVTDQAKNEQWDEENLHRDDEEYASDG 180

QY 417 AGGDDYEDNLHSPILSRQATGAEGKDI VHHGHRGSGALSMRQTLLEGEGDGVSTDIG 476
DB 181 AGGDDYEDNLHSPILSRQATGAEGKDI VHHGHRGSGALSMRQTLLEGEGDGVSTDIG 240

QY 477 QLAWKWSKEGNGRKEGGFKRVYLHQBGVPGSRGSI VSLPGGDDVFESEFVHAAALV 536
DB 241 QLAWKWSKEGNGRKEGGFKRVYLHQBGVPGSRGSI VSLPGGDDVFESEFVHAAALV 300

QY 537 SOSALFSKGLAEPMSDAAMVHPSEVAAGSRWKDLPEPGVRRALLVGVGIQLQQFAGI 596
DB 301 SOSALFSKGLAEPMSDAAMVHPSEVAAGSRWKDLPEPGVRRALLVGVGIQLQQFAGI 360

QY 597 NGVLYTTPQILEQAGVAVILSKFGLSSASILSSLTTLMLPCIGFAMLLDLSGRRF 656
DB 361 NGVLYTTPQILEQAGVAVILSKFGLSSASILSSLTTLMLPCIGFAMLLDLSGRRF 420

QY 657 LLLGTIPILIASLVILVNSNLDLGTLAHALLSTVSVIVFCCFVMGFGPIPNILCAEIF 716
DB 421 LLLGTIPILIASLVILVNSNLDLGTLAHALLSTVSVIVFCCFVMGFGPIPNILCAEIF 480

QY 717 PTRVGLCIAICATFTWIGDIIIVTYSLPVMLNAICLAGVFSIYAVVCLISFVFLK 776
DB 481 PTRVGLCIAICATFTWIGDIIIVTYSLPVMLNAICLAGVFSIYAVVCLISFVFLK 540

QY 777 TKGMPLVITEFFFAVGAQAAAKA 800
DB 541 TKGMPLVITEFFFAVGAQAAAKA 564

```

RESULT 6
US-10-051-909-37
; Sequence 37, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:

Db 716 PETKGMPELVITEFFSVGAKQA 737

RESULT 7

US-10-437-963-195955
; Sequence 195955, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 195955
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91853C.1.pep
US-10-437-963-195955

Query Match 69.5%; Score 2834; DB 16; Length 740;
Best Local Similarity 74.5%; Pred. No. 9.7e-235; Indels 6; Gaps 5;
Matches 553; Conservative 75; Mismatches 108;

Qy 56 MSGAVLVAIVASIGNLQGDWNTAIAAVLYIKKEFQLONEPTVEGLIVMSLIGATIVT 115
Db 1 MAGAVLVAIAASIGNLQGDWNTAIAAGVLYIKKEFNQSEPLIEGLIVAMSLIGATIT 60
Qy 116 TFGSPILSDSITGRPMILILSSILYFSGGLIMLWSPNVYLLARFVDGFGIGLAVTLVPLY 175
Db 61 TFGAVADSFGRPMILIASAVLYFVSGLVMLWAPNVYLLARLIDGFGIGLAVTLVPLY 120
Qy 176 ISIAPEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPDWIRIMGLVLAIPSLFFFG 235
Db 121 ISTAFTDIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPDWIRIMGLVLAIPSLIFAL 180
Qy 236 TIFYPESPRWLVSQGRMAEAKVQLKRGKDDVSGELSLLEGLVGGDTSEIYYIIGP 295
Db 181 TIFYPESPRWLVSQGRMAEAKVQLKRGREDVSGEMALLVEGLVGKDKTKEIYYIIGP 240
Qy 296 ATEAADLVTDGKEQITLYGPEEGSWIARPSKGPIMLGSVLSLASRHGSMVNSQVPLM 355
Db 241 DDELADEGLAP-DPEKIKLYGPEEGLSWARPVHGQSALGSLGSLSRHGSVMVSGKPLV 299
Qy 356 DPVITLFGSVHENMPQAGGSMRSTLPNFGSMFVSTDQHAQKNEQWDEENLHRDDEYASD 415
Db 300 DPVITLFGSVHEKMPKIMSGMRSTLPNFGSMFVAEQQAQKGDWDAES-OREGEDYGSD 358
Qy 416 GAGGYEDNLSPLSRQATGAEGKDI-VHHGRGSALSMMRQTLLGEGDGVSTDTIGG 474
Db 359 HGGDIEDLSQSPILSRQATSVGEKEIAAPHGSIINGAVG--RSSSLMQGGEAVSMGIGG 416
Qy 475 GWOLAWKSKKEGNGRKEGKFRVYLHQEVGPGSRGSIIVSLPGGQVDFGSEFVHAAA 534
Db 417 GWOLAWKWTREKADGEKGFQRIYHHEGVTGDRRGSILSLP-GGDVPPGGEFVQAAA 475
Qy 535 LVQSALFSKGLAEPRMSDAAMVHPSVAAGKSRWKDLFEPGVREALLVGVGIQLQOFA 594
Db 476 LVSPALYSKELMEQRLAGPAMVHPSQAVAKGPKWADLFEPGVKHALFVGIGIQLQOFA 535
Qy 595 GINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLITLLMLPCIGFAMLLMDLSGR 654
Db 536 GINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLITLLMLPCIGFAMLLMDLSGR 595

; APPLICANT: Allen, Steve
; APPLICANT: Helencjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-03-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 37
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-051-909-37

Query Match 69.5%; Score 2834; DB 13; Length 740;
Best Local Similarity 74.5%; Pred. No. 9.7e-235;
Matches 553; Conservative 75; Mismatches 108; Indels 6; Gaps 5;

Qy 56 MSGAVLVAIVASIGNLQGDWNTAIAAVLYIKKEFQLONEPTVEGLIVMSLIGATIVT 115
Db 1 MAGAVLVAIAASIGNLQGDWNTAIAAGVLYIKKEFNQSEPLIEGLIVAMSLIGATIT 60
Qy 116 TFGSPILSDSITGRPMILILSSILYFSGGLIMLWSPNVYLLARFVDGFGIGLAVTLVPLY 175
Db 61 TFGAVADSFGRPMILIASAVLYFVSGLVMLWAPNVYLLARLIDGFGIGLAVTLVPLY 120
Qy 176 ISIAPEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPDWIRIMGLVLAIPSLFFFG 235
Db 121 ISTAFTDIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPDWIRIMGLVLAIPSLIFAL 180
Qy 236 TIFYPESPRWLVSQGRMAEAKVQLKRGKDDVSGELSLLEGLVGGDTSEIYYIIGP 295
Db 181 TIFYPESPRWLVSQGRMAEAKVQLKRGREDVSGEMALLVEGLVGKDKTKEIYYIIGP 240
Qy 296 ATEAADLVTDGKEQITLYGPEEGSWIARPSKGPIMLGSVLSLASRHGSMVNSQVPLM 355
Db 241 DDELADEGLAP-DPEKIKLYGPEEGLSWARPVHGQSALGSLGSLSRHGSVMVSGKPLV 299
Qy 356 DPVITLFGSVHENMPQAGGSMRSTLPNFGSMFVSTDQHAQKNEQWDEENLHRDDEYASD 415
Db 300 DPVITLFGSVHEKMPKIMSGMRSTLPNFGSMFVAEQQAQKGDWDAES-OREGEDYGSD 358
Qy 416 GAGGYEDNLSPLSRQATGAEGKDI-VHHGRGSALSMMRQTLLGEGDGVSTDTIGG 474
Db 359 HGGDIEDLSQSPILSRQATSVGEKEIAAPHGSIINGAVG--RSSSLMQGGEAVSMGIGG 416
Qy 475 GWOLAWKSKKEGNGRKEGKFRVYLHQEVGPGSRGSIIVSLPGGQVDFGSEFVHAAA 534
Db 417 GWOLAWKWTREKADGEKGFQRIYHHEGVTGDRRGSILSLP-GGDVPPGGEFVQAAA 475
Qy 535 LVQSALFSKGLAEPRMSDAAMVHPSVAAGKSRWKDLFEPGVREALLVGVGIQLQOFA 594
Db 476 LVSPALYSKELMEQRLAGPAMVHPSQAVAKGPKWADLFEPGVKHALFVGIGIQLQOFA 535
Qy 595 GINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLITLLMLPCIGFAMLLMDLSGR 654
Db 536 GINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLITLLMLPCIGFAMLLMDLSGR 595
Qy 655 RELLIGTPTILASLILVWNLIDGLAHALLSTVSVIVYFCFCVFGFGFIPNILCAE 714
Db 596 RELLATIPILVALAILVNLIDVGTWVHASLSTVSVILYFCFCVFGFGFIPNILCAE 655
Qy 715 IPTTVRGICIAICFTWIGDIIIVTSLPVMNLAIGLAVPSIYAVVCLISFVFVFLKY 774
Db 656 IPTTVRGICIAICFTWIGDIIIVTSLPVMNLAIGLAVPSIYAVVCLISFVFVFLKY 715
Qy 775 PETKGMPELVITEFFSVGAKQA 796

QY 655 RPLLTGTPILIASVILVSNLIDLTALHALLSTVSVIVYFCCFVNGFGPIINILCAE 714
 Db 596 RFLLTATPILVALAAILNLVNDVGTWVHASLSTVSVILYFCCFVNGFGPIINILCAE 655
 QY 715 IPTVRVGLCIACTFWIGDIIIVTYSLPVNLNAIGLAGVFSIYAVVCLISFVFLKV 774
 Db 656 IPTTVRGICIACTFWIGDIIIVTYSLPVNLNAIGLAGVFSIYAVVCLIAFLFVPMKV 715
 QY 775 PETKGMPLVITEFFAVGAKQA 796
 Db 716 PETKGMPLVITEFFAVGAKQA 737
 RESULT 8
 US-10-051-902-8
 ; Sequence 8, Application US/10051902
 ; Publication No. US20020178468A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Hitz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB-1163
 ; CURRENT APPLICATION NUMBER: US/10/051,902
 ; PRIOR FILING DATE: 2002-01-17
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 8
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 US-10-051-902-8

Query Match 65.6%; Score 2674; DB 13; Length 737;
 Best Local Similarity 69.9%; Pred. No. 5.9e-221;
 Matches 522; Conservative 92; Mismatches 115; Indels 18; Gaps 8;
 QY 56 MSGAVLVAIVASIGNLQGDWNTAIAAVLYIKKEFQIQNEPTVEGLVSNLIGATVIT 115
 Db 1 MKGAVLVAIAASIGNFLQGDWNTAIIAGANGYIKKDLALGT--TWERLVVGNLSLIGATVIT 58
 QY 116 TFSGLSDSISGRPRMLLSSILYFFSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 175
 Db 59 TCSGFIADWLGRPRPMWIISSVLYFLGGLVMLWSPNVYVLLARLLDGFGLAVTLVPLY 118
 QY 176 ISETAPSEIRGLNLTLPQFSGSGGMFLSYCMVFGMSLSPSPDWIRMLGVLAISLFFGL 235
 Db 119 ISETAPSEIRGLNLTLPQFSGSGGMFLSYCMVFGMSLSPSPDWIRMLGVLAISLFFGL 178
 QY 236 TIFVLPESPRMLVSKGRMAEAKVQLKRGKDDVSGELSLLEGLVGGDTSTIEEYIIGP 295
 Db 179 TIFVLPESPRMLVSKGRMAEAKVQLKRGKDDVSGELSLLEGLVGGDTSTIEEYIIGP 238
 QY 296 ATEAADDLVTDGKEQITLYGPEEQSWIARPSKGPIMGLSVLSLASHSGMNVQSVPLM 355
 Db 239 ADDVADGHEHATEKDKIRLYGSQAGLSWLSKPVGTQSSIG---LASHHSGIINQSMPLM 294
 QY 356 DPVTLFGSVHENMPCAG--GSMRSTLFPNFGSMFSTVDQAKNEQWDEENLHRDDEYA 413
 Db 295 DPLVTLFGSVHENMPCAG--GSMRSTLFPNFGSMFSTVDQAKNEQWDEENLHRDDEYA 354
 QY 414 SDGAGGYEDNLHSPILSRQATGAKGDIVHGHGRGSA--SMRQTLLGEG--GDGVSSTD 471
 Db 355 SDATRGSDDDLHSPILSRQATGAKGDIVHGHGRGSA--SMRQTLLGEG--GDGVSSTD 413
 QY 472 IGGWQWLAKWSXEGENRKEGGFKVYLHOECVPGSRGSIIVSLPGGDDVPEGSEFVH 531
 Db 414 IGGWQWLAKWTDK--GEDGKQGGGFKRILYHEEGVSASRRGSIIVSLPCEG-----EFVQ 466

QY 532 AALVYSQSALFSKGLAEPMSDAAMVHPSEVAAGSRMKDLFEPCGVRRALLVGVGIQLQ 591
 Db 467 AALVYSQALYSKELIDGHPGVAMVHPSEVASKGSKALDEPGVKALVGVGIQLQ 526
 QY 592 QFAGINGVLYYTPQILEQAGAVIILSKFGLSSASASIIISLTLMLPCTGIFAMLLMDL 651
 Db 527 QFSGINGVLYYTPQILEEAGVEVLLSDIGIGSESASFILISATFTFLMLPCIGVAMKLMV 586
 QY 652 SGRFLLITPILIASVILVSNLIDLTALHALLSTVSVIVYFCCFVNGFGPIINIL 711
 Db 587 SGRRQLLTTPVILVSNLIDLTALHALLSTVSVIVYFCCFVNGFGPIINIL 646
 QY 712 CAEIPFTRVRGLCIACTFWIGDIIIVTYSLPVNLNAIGLAGVFSIYAVVCLISFVVF 771
 Db 647 CSEIPFTRVRGLCIACTFWIGDIIIVTYSLPVNLNAIGLAGVFSIYAVVCLISFVVF 706
 QY 772 LKVPETKGMPLVITEFFAVGAKQA 798
 Db 707 LKVPETKGMPLVITEFFAVGAKQA 733
 RESULT 9
 US-10-051-909-8
 ; Sequence 8, Application US/10051909
 ; Publication No. US20020199217A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Helentjaris, Tim
 ; APPLICANT: Hitz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB1163 US CIP
 ; CURRENT APPLICATION NUMBER: US/10/051,909
 ; CURRENT FILING DATE: 2002-01-17
 ; PRIOR APPLICATION NUMBER: 60/083,044
 ; PRIOR FILING DATE: April 24, 1998
 ; NUMBER OF SEQ ID NOS: 38
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 8
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 US-10-051-909-8

Query Match 65.6%; Score 2674; DB 13; Length 737;
 Best Local Similarity 69.9%; Pred. No. 5.9e-221;
 Matches 522; Conservative 92; Mismatches 115; Indels 18; Gaps 8;
 QY 56 MSGAVLVAIVASIGNLQGDWNTAIAAVLYIKKEFQIQNEPTVEGLVSNLIGATVIT 115
 Db 1 MKGAVLVAIAASIGNFLQGDWNTAIIAGANGYIKKDLALGT--TWERLVVGNLSLIGATVIT 58
 QY 116 TFSGLSDSISGRPRMLLSSILYFFSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 175
 Db 59 TCSGFIADWLGRPRPMWIISSVLYFLGGLVMLWSPNVYVLLARLLDGFGLAVTLVPLY 118
 QY 176 ISETAPSEIRGLNLTLPQFSGSGGMFLSYCMVFGMSLSPSPDWIRMLGVLAISLFFGL 235
 Db 119 ISETAPSEIRGLNLTLPQFSGSGGMFLSYCMVFGMSLSPSPDWIRMLGVLAISLFFGL 178
 QY 236 TIFVLPESPRMLVSKGRMAEAKVQLKRGKDDVSGELSLLEGLVGGDTSTIEEYIIGP 295
 Db 179 TIFVLPESPRMLVSKGRMAEAKVQLKRGKDDVSGELSLLEGLVGGDTSTIEEYIIGP 238
 QY 296 ATEAADDLVTDGKEQITLYGPEEQSWIARPSKGPIMGLSVLSLASHSGMNVQSVPLM 355
 Db 239 ADDVADGHEHATEKDKIRLYGSQAGLSWLSKPVGTQSSIG---LASHHSGIINQSMPLM 294
 QY 356 DPVTLFGSVHENMPCAG--GSMRSTLFPNFGSMFSTVDQAKNEQWDEENLHRDDEYA 413
 Db 295 DPLVTLFGSVHENMPCAG--GSMRSTLFPNFGSMFSTVDQAKNEQWDEENLHRDDEYA 354

414 SDGAGDYEDNLHSLRQATGAEGKDIIVHGHRSAL-SMRQTLLEG-GEVSSSTD 471
 355 SDATGSDDDLHSLRQATGAEGKDIIVHGHRSAL-SMRQTLLEG-GEVSSSTD 413
 472 IGGWQIAWKEKENGKRGKGFVYVHQBGVPGSRGSIIVSLPGGDDVFESEFVH 531
 414 IGGWQIAWKEKENGKRGKGFVYVHQBGVPGSRGSIIVSLPGGDDVFESEFVH 466
 532 AAALVQSALFSGKLAERPMDSAAVHPSEVAAKGSRWKDLFEPGVRRALLVGVGIQLQ 591
 467 AAALVQSALFSGKLAERPMDSAAVHPSEVAAKGSRWKDLFEPGVRRALLVGVGIQLQ 526
 592 QFAGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLLMPCIGFAMLLMDL 651
 527 QFSGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLLMPCIGFAMLLMDL 586
 652 SCRFLLIGTPIILASIVLVSNLIDLTGLAHALLSTVSVIVYFCCFVMGFGPIPNIL 711
 587 SGRQQLLTTPILASIVLVSNLIDLTGLAHALLSTVSVIVYFCCFVMGFGPIPNIL 646
 712 CAEIFPVRVGLCIACAFTFWIGDIIIVTSLPVLNAIGLAGVFSIYAVVCLISFVVF 771
 647 CSEIFPVRVGLCIACAFTFWIGDIIIVTSLPVLNAIGLAGVFSIYAVVCLISFVVF 706
 772 LKVPETKGMPLVITEFFPAVGAQAAA 798
 707 LKVPETKGMPLVITEFFPAVGAQAAA 733

RESULT 10
 US-10-051-902-29
 ; Sequence 29, Application US/10051902
 ; Publication No. US20020178468A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Hitz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB-1163
 ; CURRENT APPLICATION NUMBER: US/10/051.902
 ; PRIOR FILING DATE: 2002-01-17
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 29
 ; LENGTH: 729
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 US-10-051-902-29

Query Match 63.4%; Score 2584; DB 13; Length 729;
 Best Local Similarity 68.8%; Pred. No. 3.3e-213;
 Matches 512; Conservative 89; Mismatches 123; Indels 20; Gaps 8;
 56 MSGAVLVAIVASIGNLLQGDWNTAAAVLYIKKEFQIQNEPTVEGLIVSMISLIGATIVT 115
 1 MSGAVLVAIAAAGVNLQGDWNTAAAVLYIKKEFNESNPSVEGLIVAMSLIGATIT 60

116 TFGSPLSDSISGRPMILSLISILYFFSGILMLWSPNVYLLARFVDGFGIGLAVTLVPLY 175
 61 TCSGGVADWLGRRPMILSLISILYFFSGILMLWSPNVYLLARFVDGFGIGLAVTLVPLY 120
 176 ISEIAPSEIRGLNLTLPQFSGSGMFLSYCMVFGNLSLSPDWRIMLGVLAIPLSPFFGL 235
 121 ISEIAPSEIRGLNLTLPQFSGSGMFLSYCMVFGNLSLSPDWRIMLGVLAIPLSPFFGL 180

236 TIFVLPESRMLVSKRMAEAKVQLKRGKDDVSGELSLLEGLVGGDTSIEEYIIGP 295
 181 TVFFLPESRMLVSKRMAEAKVQLKRGKDDVSGELSLLEGLVGGDTSIEEYIIGP 240
 296 ATEAADDLVTDGKEQITLYGPEEGQSWTARPSKGPIMLGVLAIPLSPFFGL 354

241 ADEVTDHDIADVDDQIKLYGAEEGLSWARVKG---GSTMSVLSRHGSTMSSRQSL 296
 355 MDPIVTLFGSVHENPQAGGSMRSTLFPNFGSMFVSTQDQAKNEQWDEENLHRDDEEYAS 414
 297 IDPLVTLFGSVHENPQAGGSMRSTLFPNFGSMFVSTQDQAKNEQWDEENLHRDDEEYAS 355
 415 DGAGDYEDNLHSLRQATGAEGKDIIVHGHRSAL-SMRQTLLEG-GEVSSSTDIG 474
 356 D-HGDDSEDDLHSLRQATGAEGKDIIVHGHRSAL-SMRQTLLEG-GEVSSSTDIG 413
 475 GQCLAWKSEKENGKRGKGFVYVHQBGVPGSRGSIIVSLPGGDDVFESEFVHAAA 534
 414 GQVAVKWTEDSEDESGQKE-----EGFPGSRGSIIVSLPGGDDGTGE-ADFVQASA 462
 535 LVSQSALFSGKLAERPMDSAAVHPSEVAAKGSRWKDLFEPGVRRALLVGVGIQLQ 594
 463 LVSQPALYSKDLKKEHTIGPAMVHPSE-TTKGSIWHLHDPGVKRALVGVGIQLQ 521
 595 GINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLLMPCIGFAMLLMDLSGR 654
 522 GINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLLMPCIGFAMLLMDLSGR 581
 655 RFLIGTPIILASIVLVSNLIDLTGLAHALLSTVSVIVYFCCFVMGFGPIPNILCAB 714
 582 RTLLTTPILASIVLVSNLIDLTGLAHALLSTVSVIVYFCCFVMGFGPIPNILCSE 641
 715 IPTRVRVGLCIACAFTFWIGDIIIVTSLPVLNAIGLAGVFSIYAVVCLISFVVFELKY 774
 642 IPTRVRVGLCIACAFTFWIGDIIIVTSLPVLNAIGLAGVFSIYAVVCLISFVVFELKY 701

775 PETKGMPLVITEFFPAVGAQAAA 798
 702 PETKGMPLVITEFFPAVGAQAAA 725

RESULT 11
 US-10-051-909-29
 ; Sequence 29, Application US/10051909
 ; Publication No. US20020199217A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Helentjaris, Tim
 ; APPLICANT: Hitz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB1163 US CIP
 ; CURRENT APPLICATION NUMBER: US/10/051.909
 ; CURRENT FILING DATE: 2002-01-17
 ; PRIOR FILING DATE: 60/083,044
 ; PRIOR FILING DATE: April 24, 1998
 ; NUMBER OF SEQ ID NOS: 38
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 29
 ; LENGTH: 729
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 US-10-051-909-29

Query Match 63.4%; Score 2584; DB 13; Length 729;
 Best Local Similarity 68.8%; Pred. No. 3.3e-213;
 Matches 512; Conservative 89; Mismatches 123; Indels 20; Gaps 8;
 56 MSGAVLVAIVASIGNLLQGDWNTAAAVLYIKKEFQIQNEPTVEGLIVSMISLIGATIVT 115
 1 MSGAVLVAIAAAGVNLQGDWNTAAAVLYIKKEFNESNPSVEGLIVAMSLIGATIT 60

116 TFGSPLSDSISGRPMILSLISILYFFSGILMLWSPNVYLLARFVDGFGIGLAVTLVPLY 175
 61 TCSGGVADWLGRRPMILSLISILYFFSGILMLWSPNVYLLARFVDGFGIGLAVTLVPLY 120
 176 ISEIAPSEIRGLNLTLPQFSGSGMFLSYCMVFGNLSLSPDWRIMLGVLAIPLSPFFGL 235
 121 ISEIAPSEIRGLNLTLPQFSGSGMFLSYCMVFGNLSLSPDWRIMLGVLAIPLSPFFGL 180

Db 121 ISETAPPETIRGLNLTLPQFTSGGNGFLSYCMVFGMSLMPSPSRWMLMLGVLPISLVFFFL 180
Qy 236 TIFYLPEPRWLVSKGRMAEAKVLOKRGKDDVSGLSLLLEGVGGDTSIEEYIIGP 295
Db 181 TVFFLPESPRMLVSKGRMLEAKRVQRURGREDSVGENALLVEGLIGGETTIEEYIIGP 240
Qy 296 ATEAADDLVTDGKEQITLYGPEEQSWIARPSKPIMLGSLVLSLASHGHS-VYNQSVPL 354
Db 241 ADEVTDHDAVDKQIKLYGAEEGLSWARPVKG---GSTMSVSRHSGTMSRRQSL 296
Qy 355 MDPIVTLFGSVHNNPQAGSMRSTLFPNFGSMFVSVDQHAQKNEQWBNLHRDDEYAS 414
Db 297 IDPLVTLFGSVHNNPQAGSMRSTLFPNFGSMFVSVDQHAQKNEQWBNLHRDDEYAS 355
Qy 415 DGAGDYEDNLHSPILSRQATCAEGKQIVHHRGHSALSMRQTLLEGSGDVSSTDIGG 474
Db 356 D-HGDDSEDDLHSPILSRQATCAEGKQIVHHRGHSALSMRQTLLEGSGDVSSTDIGG 413
Qy 475 QHQLAWKSEKGEKNGRKEGKRVYVHQBEGVPSRRGSIIVSLPGGGDVFEQSEFVHAAA 534
Db 414 QHQAQWTEREDESQKE-----EGFPGSRGSIIVSLPGGGDGTGE-ADFVQASA 462
Qy 535 LVSQALSKGLAEPMSDAAWHPSEVAAKGSRWKOLFEPGVRRALLVGVGIQILOQFA 594
Db 463 LVSQALSKGLAEPMSDAAWHPSEVAAKGSRWKOLFEPGVRRALLVGVGIQILOQFA 521
Qy 595 GINGVLYTTPQLEQAGVAVILSKFGJSSASASILISLTLLMLPCIGFAMLLMDLSGR 654
Db 522 GINGVLYTTPQLEQAGVAVILSKFGJSSASASILISLTLLMLPCIGFAMLLMDLSGR 581
Qy 655 RFLGTTPIILASIVAVLSNLDLGLTLAHLSTVSIVVYFCFVNGGPIPNILCAE 714
Db 582 RFLGTTPIILASIVAVLSNLDLGLTLAHLSTVSIVVYFCFVNGGPIPNILCAE 641
Qy 715 IFTVRVGLCIAICAFITWIGDIIVTYSPLVYMLNAIGLAGVFSIYAVVCLISFVFLKV 774
Db 642 IFTVRVGLCIAICAFITWIGDIIVTYSPLVYMLNAIGLAGVFSIYAVVCLISFVFLKV 701
Qy 775 PETKGMPLVITERFAVCAKQAAA 798
Db 702 PETKGMPLVITERFAVCAKQAAA 725

RESULT 12

US-10-424-599-153166
; Sequence 153166, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J.
; APPLICANT: Kovalic David K.
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 153166
; LENGTH: 742
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(742)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_WRT3847_109333C.1.pap
US-10-424-599-153166

Query Match 63.3%; Score 2581.5; DB 15; Length 742;
Best Local Similarity 67.3%; Pred. No. 5.5e-213;

Matches 511; Conservative 101; Mismatches 110; Indels 37; Gaps 13;
Qy 56 MSGAVTVAIVASIGNLQGNDAATAAAYLYIKKSFQIQNEPTVEGLIVSMGLIGATVIT 115
Db 1 MKGTULVAIAAIGNILQGNDAATAGAIYVYIKDALQCT--TMGLVVAIVSGLIGATVIT 58
Qy 116 TSPGUSDISIGRRPMLISSLILYFFSGLIMLSPNVYVLLARFVDFGIGIAGLAVTLVPL 175
Db 59 TCSGPTADMLGRRPMMIISVLYFLGGLVWMLSPNVYVLLARLDDGFGIAGLAVTLVPL 118
Qy 176 ISEIAPSEIRGLNLTLPQFSGGGMFLSYCMVFGMSLSPSPDRIMLVIAIPLSPFFGL 235
Db 119 ISETAPSEIRGLNLTLPQFSGGGMFLSYCMVFGMSLSPSPDRIMLVIAIPLSPFFGL 178
Qy 236 TIFYLPEPRWLVSKGRMAEAKVVL-----QKLRGKDDVSGLSLLLEG-LEYGGDTS 287
Db 179 TIFFLPESS---SSGSVSKXNMLRROTCKSORLURGREDSVSGEMTLLVEGSTXIGDTS 234
Qy 288 IBEYIIGRATEADDLVTDGKEQITLYGPEEQSWIARPSKPIMLGSLVLSLASHGSM 347
Db 235 IBEYIIGXADQVADGHEHATEKDIRLYGQAGLSWAKPVTGQSSIG----LASRHGSI 290
Qy 348 VNQSVPLMDPIVTLFGSVHNNPQ--AGSMRSTLFPNFGSMFVSVDQHAQKNEQWBNL 405
Db 291 INQSVPLMDPIVTLFGSITHEKLPETGAGSMESTLFPNFGSMFSTAEPAKNEQWBNL 350
Qy 406 HRDDREYASDAGGDIYENLHSPILSRQATCAEGKDI---VHGHRSAL-SMRQTL 460
Db 351 OREGEDYMSDAAAGSDDDLHSPILSRQATCAEGKDI---VHGHRSAL-SMRQTL 406
Qy 461 GSG-GDGVSTDIGGWLAWKSEKGEKRVYVHQBEGVPSRRGSIIVSLPG 519
Db 407 MGSGEQGSGTIGGWLAWKWTDX-GEQKQGGFRIVLHBEQVSAASRGSIIVSLPG 465
Qy 520 GGDVFEQSEFVHAAALVQSALFSKGLAEPMSDAAWHPSEVAAKGSRWKDLFEPGVRR 579
Db 466 EG-----EFVQAAALVSPALYSKELIDGHPVGPAMVHPSETASKGPSKALLFEPGVRR 519
Qy 580 ALLVGVGIQILOQFAGINGVLYTTPQLEQAGVAVILSKFGJSSASASILISLTLLML 639
Db 520 ALLVGVGIQILOQFAGINGVLYTTPQLEQAGVAVILSKFGJSSASASILISLTLLML 579
Qy 640 PCIGFAMLLMDLSGRFLLGTIPILIASLVLSNLDLGLTLAHLSTVSIVVYFCF 699
Db 580 PCIGFAMLLMDLSGRFLLGTIPILIASLVLSNLDLGLTLAHLSTVSIVVYFCF 639
Qy 700 FVNGGPIPNILCAEIPFTRVGLCIAICAFITWIGDIIVTYSPLVYMLNAIGLAGVFSIY 759
Db 640 FVNGGPIPNILCAEIPFTRVGLCIAICAFITWIGDIIVTYSPLVYMLNAIGLAGVFSIY 699
Qy 760 AVVCLISFVFLKVETKGMPLVITERFAVCAKQAAA 798
Db 700 AVVCLISFVFLKVETKGMPLVITERFAVCAKQAAA 738

RESULT 13

US-10-425-114-47258
; Sequence 47258, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 47258

```

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 251277
; LENGTH: 742
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(742)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_68931C.1.pep
; US-10-424-599-251277

Query Match 61.3%; Score 2496.5; DB 15; Length 742;
Best Local Similarity 65.8%; Pred No. 1.2e-205;
Matches 505; Conservative 84; Mismatches 121; Indels 57; Gaps 12;

QY 56 MSGAVLVAIVASIGNLLQGDWNTAT-----AAAVLIKKEP 91
DB 1 MSGAVLVAIVAAIGNLLQGDWNTATAGHIFIRLANXNDIOPSLIGFLCVAGSILYIKKEP 60
QY 92 QONERTVEGLVSKSLGATVITVTFSGPLSDSIGRRPMLILSSILYFSGILMLWSPNV 151
DB 61 NQOSEPTIEGLVAVSLGATVITVTCGSLDGLRRPMLIISSILYFVGLVMLWSPNV 120
QY 152 YVLLARFVDFGIGLAVTLVPLVYISETAPSEIRGLNLTLPQFSGGGMFLSYCMVFGMS 211
DB 121 YILLFARLLDGLGIGLAVTLVPLVYISETAPSEIRGLNLTLPQFSGAGMFFSYCMVFAMS 180
QY 212 LSPSPDWRLMGLVLAIPSLFFGLTIFVLPSPRMLVSKGRMAEAKVILKLRGDDVSG 271
DB 181 LTKAPNWRMLGLVLSIPSLIYFALTLPFSPRMLV-KGRMLEAKVILKLRGDDVAG 239
QY 272 ELSLLEGEVGGDTSIEEYIIPATEAADLVTGDKKEQITLYGPEEQSWIARSKGP 331
DB 240 EWALLVEGLVGRDIAIEEYIIGPAXEFS-----EAEQIKLYGTAEVSWIAKPVTCQ 292
QY 332 IMLGSVLSLASRHGSMVNSQVPLMDPIVTLFGSVHENMPQAGSMRSTLPNFGSMFSVT 391
DB 293 SSIG-----LVSRKGSMAQS-GLVDPVLKLVGSHVHEKLPETG-----STLPHFGSMFSVG 343
QY 392 DQAKNEQWDEENLHDDDEYASDGAG-GDYEDNLHSLPLSRQATGAEGKDVHGHGRS 450
DB 344 GNQPRNEDWDESIAREGDDYSDAVDITDDSDNLQSPILSRQATSAR-RDMPAPAQG-- 400
QY 451 ALSMRGQTLIGEGDGVSTDIGGQWLAKWSEKEGNGRKEGGFKRYLHQEGVPGSR 510
DB 401 --SMRQGSLL--QGEPAAGNSGIGGQWLAKWSETEGV-----FKRYLHQEGGPGSR 449
QY 511 RGSIVSLPG-GGDVFESEFVHAALVQSALFSKGLAEPRMSDAAMVHPSEVAAGSRW 569
DB 450 RGSIIISLPGCDAPTLTGDGIVQAALVQSALYNKELMHQPVGPAMIHPSQTAAGPSW 509
QY 570 KDLFEQVRRALLVGVGIQILOQFAGINGVLYVYTPQILEQAGVAVILSKFGLSSASIL 629
DB 510 SDLLEFGVHALVGVGIQILOQFSGINGVLYVYTPQILEQAGVGVYLLNLGLSTGASPL 569
QY 630 ISSLTLLMLPCIGFAMLLMDLSRRFLLLGTIPILIASLIVLVSNLIDLTGLAHALLS 589
DB 570 ISSVTTLLMLPCIAVANRLMDISGRRTLLTTPVILVSLILVIGLSVLDSTINAFIS 629
QY 690 TVSVIVYFCCFVMGFGPIPNILCAETFPTRVRLGICIAICAFTFWIGDITVYSLPVLNA 749
DB 630 TSSVIVYFCCFVMGFGPIPNILCSEFPTRVRLGICIAICAFTFWICDITVYSLPVLNS 689
QY 750 IGLGVFSIYAVVCLISFVFLKVPETKGMPLVETEFPAVGAQQA 796
DB 690 VGLGVFGMVAVVCIIAWVFLKVPETKGMPLVETEFPAVGAQQA 736

```

```

; LENGTH: 728
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700728773_FLI.pep
; US-10-425-114-47258

Query Match 62.4%; Score 2543; DB 15; Length 728;
Best Local Similarity 68.4%; Pred. No. 1.1e-209;
Matches 508; Conservative 83; Mismatches 120; Indels 32; Gaps 10;

QY 56 MSGAVLVAIVASIGNLLQGDWNTATIAAAVLIYKKEFOLONEPTVEGLIVSMISLIGATVIT 115
DB 10 MSGAVLVAIVAAIGNLLQGDWNTATAGSILYIKKEFNLOQSEPTIEGLIVAMSLIGATVIT 69
QY 116 TFSGPLSDSIGRRPMLILSSILYFSGILMLWSPNVYVLLARFVDFGIGLAVTLVPLV 175
DB 70 TCSGPLSDLGLRRPMLIISSILYFSGILVMLWSPNVILFARLLDGLGIGLAVTLVPLV 129
QY 176 ISETAPSEIRGLNLTLPQFSGGGMFLSYCMVFGMSLSPSPDWRLMGLVLAIPSLFFGL 235
DB 130 ISETAPSEIRGLNLTLPQFSGAGMFFSYCMVFASILTAKPNWRMLGLVLSIPSLIYFAL 189
QY 236 TIFVLPSPRMLVSKGRMAEAKVILKLRGDDVSGLSLLEGLVGGDTSIEEYIIGP 295
DB 190 TLFFLPSPRMLVSKGRMLEAKVILKLRGDDVAGVALLVEGLVGGDTSIEEYIISP 249
QY 296 ATEAADLVTGDKKEQITLYGPEEQSWIARSKGPIMLGSVLSLARHSGSMVNSQVPLM 355
DB 250 ANEFS-----DAEQIKLYGTAEVSWIAKPVTCQSSIG-----LVSRKGSMAQSALV 297
QY 356 DPIVTLFGSVHENMPQAGSMRSTLPNFGSMFSVTQHAKNEQWDEENLHRRDDEYASD 415
DB 298 DPLVKLFSGVHEKLPETG-----STLPHFGSMFSVGGNQPRNEDWDESIAREGDDTVSD 353
QY 416 GAG-GDYEDNLHSLPLSRQATGAEGKDVHGHGRSALSRRQTLLEGEGDGVSTDIG 474
DB 354 AADTDDSDNLQSPILSRQATSAR-RDMPAPAQG---SMRQGSLL--QGEPAAGNSGIGG 406
QY 475 GWQLAKWSEKEGNGRKEGGFKRYLHQEGVPGSRGSIIVSLPG-GGDVFESEFVHA 533
DB 407 GWQLAKWSETEGV-----FKRYLHQEGGPGSRGSIISLPGCDAPTLTGDGEIVQAA 459
QY 534 ALVSQSALFSKGLAEPRMSDAAMVHPSEVAAGSRKWKDLPPGVRRALLVGVGIQILOQF 593
DB 460 ALVSQSALYNKELMHQPVGPAMIHPSQTAAGPSWSDLLPEPGVGHVHALVGVGIQILOQF 519
QY 594 AGINGVLYVYTPQILEQAGVAVILSKFGLSSASILISSLTLLMLPCIGFAMLLMDLSG 653
DB 520 SGINGVLYVYTPQILEQAGVGVYLLNLGLSTGASFLISSVTTLLMLPCIAVANRLMDISG 579
QY 654 RFLILGTIPILIASLIVLVSNLIDLTGLAHALLSTVSVIVYFCCFVMGFGPIPNILCA 713
DB 580 RPPLLLTIPVLIVSLILVIGSVLVDSTINAFISTSSVIVYFCCFVMGFGPIPNILCS 639
QY 714 BIFPTRVRLGICIAICAFTFWIGDITVYSLPVLNAIGLAGVFSIYAVVCLISFVFLK 773
DB 640 BIFPTRVRLGICIAICAFTFWICDITVYSLPVLNSVGLGVFGMVAVVCIIAWVFLK 699
QY 774 VPETKGMPLVETEFPAVGAQQA 796
DB 700 VPETKGMPLVETEFPAVGAQQA 722

```

```

RESULT 14
US-10-424-599-251277
; Sequence 251277, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei

```

Db 595 TLPVMLNVVGLAGVGVAVVAVCVLALAFVFIKVPETKGMPLVITEFFSVGAKQA 649

Search completed: October 13, 2004, 11:54:38
Job time : 305.735 secs

```
RESULT 15
US-10-425-114-37283
; Sequence 37283, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 37283
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB83-005-C10_FLI pep
US-10-425-114-37283

Query Match 59.9%; Score 2442.5; DB 15; Length 661;
Best Local Similarity 72.1%; Pred. No. 4.4e-201;
Matches 472; Conservative 76; Mismatches 100; Indels 7; Gaps 6;

QY 143 LLMWSPNVVYLLARFVDFGICGLATVLPYLIYSEIAPSEIRCLNTLPQPSGSGMFL 202
Db 1 LVMWAPSVYLLARLDGIGGLAVTLVPLYISETAPTDIRGLNTLPQPSGSGMFL 60
QY 203 SYCNVFGMSLSPDWMIGVLAIPSLFFFGITIFYPSPRWLVSKGMAEAKVQLQ 262
Db 61 SYCNVFGMSLMPKPDWMLGLVLSIPSLYFGLTVFVLPSPRWLVSKGMAEAKVQLQ 120
QY 263 LRKDDVSGEUSLLEGLVGGDTSEIYIIGPATEAADLVTDGKQITLYGPREGQS 322
Db 121 LRGREVSGENALLVEGLVGKDTRIEYIIGPDDELADGLAP-DPEKIKLYGPEGLS 179
QY 323 WIAPSKGPIMLGSLVLSASHGSM-VNQSVPMDPIVTLFGSVHENMPOAGGSMRSTLF 381
Db 180 WVARPVGQSLGSLGALGLISHGSMASQKPLVDPNVTFLGSHVHKMPIMGSMRSTLF 239
QY 382 PNFGSMFVTDQAHKNEQWDEENLHRDDEYASDAGGDEYEDNLHSPLLSRQATGAGKD 441
Db 240 PNFGSMFVADQQVKADWDAES-QREGEDYASDHGGDDIEDNLQSPILSRQATSVGEKE 298
QY 442 IVHHGHSALSMEKRTLLGEGDGVSTDIGGWQLAKWSEKENGKGGPKRVYL 501
Db 299 IA--APRGSILGAVGRSSQGGGAVSMGIGGGWQLAKWTEREGDQGGGQFRIYL 356
QY 502 HQEGVPSRRGSIYSLPGGDFVFGSEFVHAALVSQSALFSKGLAEPKMSDAAMVHPSE 561
Db 357 HEEGVQGN-RGSILSLP-GGDVPPGGEFIOAALVSQALYSKELLEQRAAGPAMVHPSE 414
QY 562 VAAKGSRWKDLFEGVPRALLVGVIGIIOOPAGINGVLYTTPQILEQAGVAVILSKFGL 621
Db 415 AVTKGPRWDLFEGVPRALLVGVIGIIOOPAGINGVLYTTPQILEQAGVAVILSKFGL 474
QY 622 SSASASILISLTLTLLMLPCTGFAMLMDSGRFLLLTGPIPIIASLVILVVSNIIDL 681
Db 475 NASSASILISALTLLMLPSIGIAMLMDSGRFLLLTATIPVLIALLVIVSNIVDVG 534
QY 682 TLHALLSTSVIVYFCFVWGFGPIINILCAIFPFRVGLGICACAFWFGDIIVTY 741
Db 535 DVAAHALSTASVIVYFCFVWGFGFVENILCAIFPTTVRGVCIAICALAFWVGDIIVTY 594
QY 742 SLPVMLNAIGLAGVFSIYAVVCLISFVFLKVPETKGMPLVITEFFFAVGAQQA 796
Db 742 SLPVMLNAIGLAGVFSIYAVVCLISFVFLKVPETKGMPLVITEFFFAVGAQQA 796
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: October 13, 2004, 11:39:29 ; Search time 9234 seconds
(without alignments)
4097.007 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSSGSLAVQTFPTDLDLR.....PLEVITEFAVGAQAQAAKA 800

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool_h/US10051909/runat_13102004_123336_19906/app_query.fasta_1.1678
-DB=GenEmbl -QFMT=fastap -SUFFIX=age -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORW=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051909 @CGN 1 1 9091 @runat_13102004_123336_19906 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.cm.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3517	86.3	2824	6	AR208565 Sequence
2	3395	83.3	2934	8	AK120560 Oryza sat
3	3155	77.4	2665	8	AY165599 Saccharum
4	3097	76.0	2378	8	AK099716 Oryza sat

5	3052.5	74.9	136267	8	AP005756 Oryza sat
6	2825.5	69.3	2800	8	AK102640 Oryza sat
7	2808	68.9	2614	8	HVU534445 Hordeum v
8	2674	65.6	2601	6	AR208568 Sequence
9	2657	65.2	2570	8	AY094465 Arabidops
10	2584	63.4	2190	6	AX506620 Sequence
11	2584	63.4	2190	8	ATH532570 Arabidops
12	2583	63.4	2368	8	AK065191 Oryza sat
13	2514	61.7	142114	8	AC073166 Oryza sat
14	2514	61.7	300957	8	AE017116 Oryza sat
15	2502	61.4	86710	8	ATF23E12 Arabidops
16	2502	61.4	137859	8	ATCHRIV83 Arabidops
17	2368	58.1	135583	8	AC136843 Oryza sat
18	2275	55.8	2205	6	AX412656 Sequence
19	2275	55.8	2205	6	AX507559 Sequence
20	2245	55.1	2190	8	ATH532571 Arabidops
21	2217	54.4	2426	8	ATSUGTRPR A.thaliana
22	2191.5	53.8	2515	8	HVU534445 Hordeum v
23	2144	52.6	103192	8	AC007369 Arabidops
24	2144	52.6	143879	8	AC069251 Gennomic s
25	2128	52.2	114918	8	AF004945 Lotus cor
26	2117	52.0	94349	8	ATF26013 Arabidops
27	2093.5	51.4	100900	2	AC121239 Medicago
28	1968	48.3	120743	8	AP004082 Oryza sat
29	1944.5	47.7	115134	8	AC126786 Medicago
30	1918	47.1	121112	8	AC144482 Medicago
31	1670.5	41.0	1692	6	AR208569 Sequence
32	1557.5	38.2	87592	8	AC099739 Oryza sat
33	1557.5	38.2	127098	8	AC144426 Oryza sat
34	1557.5	38.2	154128	8	AP000615 Oryza sat
35	1470.5	36.1	1564	8	AK067391 Oryza sat
36	1436	35.2	2661	8	AK120270 Oryza sat
37	1424	34.9	1487	6	AR208571 Sequence
38	1411	34.6	122809	8	AC127019 Medicago
39	1409	34.6	2516	8	AK068224 Oryza sat
40	1406	34.5	1625	8	AK059705 Oryza sat
41	1174	28.8	1315	8	BT009593 Triticum
42	1159.5	28.5	70311	8	AF128457 Oryza sat
43	1140	28.0	77605	8	AF119222 Oryza sat
44	1140	28.0	142852	8	AF161269 Oryza sat
45	973	23.9	106246	8	AF528565 Zea mays

ALIGNMENTS

RESULT 1	AR208565	Sequence 1 from patent US 6383776.	2824 bp	DNA	linear	PAT 20-JUN-2002
LOCUS	AR208565	Sequence 1 from patent US 6383776.	2824 bp	DNA	linear	PAT 20-JUN-2002
DEFINITION	AR208565	Sequence 1 from patent US 6383776.	2824 bp	DNA	linear	PAT 20-JUN-2002
ACCESSION	AR208565	Sequence 1 from patent US 6383776.	2824 bp	DNA	linear	PAT 20-JUN-2002
VERSION	AR208565.1	GI:21509752	2824 bp	DNA	linear	PAT 20-JUN-2002
KEYWORDS	AR208565.1	GI:21509752	2824 bp	DNA	linear	PAT 20-JUN-2002
SOURCE	Unknown.	Unknown.	2824 bp	DNA	linear	PAT 20-JUN-2002
ORGANISM	Unknown.	Unknown.	2824 bp	DNA	linear	PAT 20-JUN-2002
REFERENCE	1 (bases 1 to 2824)	Unknown.	2824 bp	DNA	linear	PAT 20-JUN-2002
AUTHORS	Allen, S.M., Hitz, W.D., Kinney, A.J. and Tingey, S.V.	Unknown.	2824 bp	DNA	linear	PAT 20-JUN-2002
TITLE	Plant sugar transport proteins	Unknown.	2824 bp	DNA	linear	PAT 20-JUN-2002
JOURNAL	Patent: US 6383776-A, 1 07-MAY-2002;	Unknown.	2824 bp	DNA	linear	PAT 20-JUN-2002
FEATURES	Location/Qualifiers	Unknown.	2824 bp	DNA	linear	PAT 20-JUN-2002
	source	Unknown.	2824 bp	DNA	linear	PAT 20-JUN-2002
	1..2824	/organism="unknown"	2824 bp	DNA	linear	PAT 20-JUN-2002
	/mol_type="unassigned DNA"	Unknown.	2824 bp	DNA	linear	PAT 20-JUN-2002
ORIGIN	1..2824	/organism="unknown"	2824 bp	DNA	linear	PAT 20-JUN-2002
	/mol_type="unassigned DNA"	Unknown.	2824 bp	DNA	linear	PAT 20-JUN-2002

Alignment Scores:	3.68e-225	Length:	2824
Pred. No.:	3517.00	Matches:	684
Score:	95.18%	Conservative:	27
Percent Similarity:	91.57%	Mismatches:	34
Best Local Similarity:	86.31%	Indels:	2
Query Match:	6	Gaps:	2
DB:			

US-10-051-909-32 (1-800) x AR208565 (1-2824)

Qy 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
Db 238 ATGGGGGCGCCGTGATGTCGCATCGCGCCCTCTATCGGCAACTTGTGCGAGGCGCTG 297
Qy 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleIleGlyGlnLeuGlnAsn 95
Db 298 GACAAATCGCAATTCGTGAGCGCGTCTGTATATAAGAGAAATCAACCTGCGAGAGC 357
Qy 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db 358 GAGCCTCTGATCAAGGCGCTCATCTGCGCATGTTCTCATTTGGGCGACAGTCATCACA 417
Qy 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgProMetLeuIleLeuSer 135
Db 418 ACATCTCCGGGCGCAAGGCGTGAATGCTGTGTAGGAGGCCATGCTGTGCGCTCGGCT 477
Qy 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu 155
Db 478 GTCTCTACTTCTGAGTGGCTGGTATGCTTTGGCGCCCAATTGTGTACATCTGTCTC 537
Qy 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
Db 538 CTCGCAAGGCTCATGATGGGTTGCGTATCGGTTTGGCGGTCCACACTTGTCTCTAC 597
Qy 176 IleSerGluIleAlaProSerGluIleArg--GlyLeuLeuAsnThrLeuProGlnPhe 194
Db 598 ATCTCCGAAACTGCACCGCACAGANATCTTGGGGCTGNTNGAACACAGTTCGCGAGTTC 657
Qy 195 SerGly---SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213
Db 658 ATGGGGTCAGAGGAGATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 717
Qy 214 ProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233
Db 718 CCCAAACCTGATGGAGGCTCATGCTGTGGATGTTCTGTCATCCCGTCACTTATNACTTT 777
Qy 234 GlyLeuThrIlePheTyrLeuProGluSerProArgTyrLeuValSerIleGlyArgMet 253
Db 778 GGAATGACTGCT 837
Qy 254 AlaGluAlaIleGlyValLeuGlnIleGlyLeuArgGlyLysAspValSerGlyGluLeu 273
Db 838 GCGAGGCGGAGAGAGTGTCAAGGCTCGGGGAGAGAGATGCTCTCANGGGGAGAG 897
Qy 274 SerLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIle 293
Db 898 GCTCTCTAGTGAAGTTTGGGGTGGTGAAGTACACGATTTTNNAGATACATCAIT 957
Qy 294 GlyProAlaThrGluAlaAlaAspLeuValThrAspGlyAspGlyGluGlnIleThr 313
Db 958 GGACCTGCCACGAGCAGCCGATGATCTGTAACTGACGGTGATAAGGAAACAAATCACA 1017
Qy 314 LeuTyrGlyProGluGlyGlnSerTrpIleAlaArgProSerIleGlyProIleMet 333
Db 1018 CTTTATGGGCGCTGAAGAGCGGAGTATGATGTCGACCTTCTAAGGGGCCCATCATG 1077
Qy 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
Db 1078 CTTTGAAGTGTGCTTCT 1137
Qy 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
Db 1138 CTTATGGATCCGATGTGACACTTTTGGTAGTGTCCATGAGAAATATGCTCAAGCTGGA 1197
Qy 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db 1198 GGAAGTATGAGGAGCAGCATGTTTCCAACTTTGGNAGTATGTTTCAAGTGTACAGATCAG 1257
Qy 394 HisAlaGlyAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413
Db 1258 CATGCCAAAATGAGCAGTGGATGAAGAGAAATCTTATAGGATGACGAGGATGACGCA 1317

Qy 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
Db 1318 TCTGATGGTGAGGAGGTGACTATGAGCAAAATCTCCATAGCCCATTTGCTGTCCAGGAG 1377
Qy 434 AlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGlySerAlaLeuSer 453
Db 1378 GCAACAGGTGCGAAGGAGGAGCAATTTGACCATGTTCCATGGAAGTGTGTTGAGC 1437
Qy 454 MetArgArgGlnThrLeuLeuGlyGluGlyValSerSerThrAspIleGly 473
Db 1438 ATGAGAGGCAACCTCTTAGGGAGGGTGGAGATGTTGTGAGCAGCACTGATATCGGT 1497
Qy 474 GlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlnAsnGlyArgLysGlu 493
Db 1498 GGGGATGGCAGCTTGTCTGGAATGTCAGAGAAGAGTGAATGTTAGAAAGGAA 1557
Qy 494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgGlySer 513
Db 1558 GGTGGTTTCAAAGAGTCTACTTGCACCAAGAGGGAGTCTCTGGCTCAAGAGGGGCTCA 1617
Qy 514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533
Db 1618 ATTGTTTCACTTCCCGGTGGTGGATGTTCTTGAAGGAGTGTGAGTGTGATGCTGCT 1677
Qy 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553
Db 1678 GCTTTAGTAGTCACTCAGCAGCATTTTCTCAAGGGTCTTGTGAACCAAGCATGTGAT 1737
Qy 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe 573
Db 1738 GCTGCCATGTTTCACTTCACTGAGTAGTCTGCAAGAGTTCAGTTTGAAGAGTGTGTT 1797
Qy 574 GluProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPhe 593
Db 1798 GAACCTGGAGTGGGCGTCCCTGTGTAGTGGGTGGTGGAAATCAGATCTTCAACAGTTT 1857
Qy 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluAlaGlyValAla 613
Db 1858 GCTGGAATAAACGGTGTCTGTACTATACCCCAAAATCTTGGAGCAAGTGTGTGGCA 1917
Qy 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuLeuSerSerLeu 633
Db 1918 GTTATTTCTTCCAAATTTGCTCTCAGCTGGCATCAGCATCTTGTATGATCTCTCTC 1977
Qy 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
Db 1978 ACTACCTTACTATGCTTCTCTGATGCTTGGCATGCTGCTTATGATCTTTCGGA 2037
Qy 654 ArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
Db 2038 AGAAGGTTTTTGTGCTAGGCAAAATTTCAATCTTGTATGATCTCTAGTATCTCTGGT 2097
Qy 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
Db 2098 GTGTCCAACTTAATGATTTGGTACATAGCCCATGCTTTGTCTTCCACCATCAGTGT 2157
Qy 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
Db 2158 ATCGTCTACTCTGCTGCTTCTGTTATGGATTTGGTCCCATCCCAACATTTATGTGCA 2217
Qy 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
Db 2218 GAGATCTTTTCCAAACAGGTTCTGTGCTCTGTATGGCATTTGTGCTTATCATCTCG 2277
Qy 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
Db 2278 ATCGGAGATATCATCTGCTACCTACACCTTCTCTGTATGCTGAATGCTATGAGCTGGC 2337
Qy 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773
Db 2338 GGTGTTTTCAGCATATGATGAGTCGATGCTTGTGATTTCTTTGTGCTCTCTCTTAAG 2397

QY 774 valProGluThrIysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
Db 2398 GTCCCTGAGACAAAGGGATGCCCTTGAGGTATTACCGAATCTTTCAGTGTGTGGG 2457

QY 794 LysGlnAlaAlaAlaLysAla 800
Db 2458 AAGCAGCGGCTGCARAGCC 2478

RESULT 2
AKI20560
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) cDNA clone:J013132C18, full
insert sequence.
ACCESSION
AKI20560.1 GI:37990183
VERSION
FJ1_CDNA; CAP trapper.
KEYWORDS
Oryza sativa (japonica cultivar-group)
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE
1
The Rice Full-length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-length cDNA Project Team:
Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group: Ohtsuki,K., Murakami,K.,
Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
Kudo,Y., Sugano,S., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN,
Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
Hara,A., Hashizume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Yoshino,M. and Hayashizaki,Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice

JOURNAL
MEDLINE
PUBMED
12869764
REFERENCE
2
AUTHORS
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Kagawa,I.,
Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
Kagawa,S., Kato,H., Kawagashira,N., Kawai,J., Kawamata,M.,
Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,
Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,
Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakahama,Y.,
Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K.,
Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M.,
Ooka,H., Osato,N., Ota,Y., Ohtsuki,K., Ohtsuki,H., Sasaki,C.,
Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K.,
Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y.,
Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M.,
Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akai,H.,
Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K.,
Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J.,
Yokomizo,S. and Yoshimura,A.
Collection, mapping, and annotation of 28K full-length cDNA clones
from japonica rice

TITLE
Unpublished
3 (bases 1 to 2934)
REFERENCE
AUTHORS
Kikuchi,S.
TITLE
Direct Submission
Submitted (31-JAN-2003) Shoshi Kikuchi, National Institute of

COMMENT

Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression: 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:shikuchi@nias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 32K full-length cDNA clones from japonica
rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/

NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K.,
Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T.,
Yamamoto,M. and Nakahama,Y.
FAIS Genome Sequencing & Analysis Group: Ohtsuki,K., Iida,Y.,
Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M.,
Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M.,
Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M.,
Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M.,
Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,
Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N.,
Ota,Y., Saitoh,H., Sakai,C., Sakai,K.,
Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K.,
Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M.,
Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Tagami,K.,
Takaku-Akai,H., S., Tanaka,T., Tomaru,A., Toya,T., Waki,K.,
Yasunishi,A. and Hayashizaki,Y.
Location/Qualifiers
1. 2934
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J013132C18"

FEATURES

source

ORIGIN

Alignment Scores:
Pred. No.: 5,45e-217 Length: 2934
Score: 3395.00 Matches: 676
Percent Similarity: 86.84% Conservative: 43
Best Local Similarity: 81.64% Mismatches: 76
Query Match: 83.31% Indels: 34
DB: 8 Gaps: 5

US-10-051-909-32 (1-800) x AKI20560 (1-2934)

QY 3 SerGlySerTrpLeuAlaValGlnThrProPheThrProAspLeuAspArgGluArg 22
Db 38 ACTCGTGGGTGGCTCTTGGCGGCGCGCGC---GCCCGCGCACCGCGTGGCGGT 94
QY 23 LeuLeuProSerVal-----ValLeuAla 30
Db 95 GCGCTGTCTGTACCCGCCACGAGAGAGATCGATCGCCCGCGCGCGGTCTCTGCT 154
QY 31 Leu-----ProGlyProLeuProProAlaSerCysSerSerGlnGluPro----- 45
Db 155 CTTCTCTCTCTCGGCTGGCGGCGGACAGGCGGATCTTGGCGGCGCTT-CCTCCAGGGAT 213
QY 46 -----ValThrSerAsp-----Asp 50
Db 214 CTCGTCTCGCGCGCGCGCGCTCGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 273
QY 51 IleLeuGluAspLysMetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsn 70
Db 274 CGGAGAGGGGTGAAGATGTCTGGGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 333
QY 71 LeuLeuGlnGlyTrpAspAsnAlaThrIleAlaAlaValLeuTyIleLysLysGlu 90

Db 334 CTGCTGAGGATGGGCAATGCCACCAATGCGAGGTGCTACTATACATAAAGAGGAA 393
Qy 91 PheGlnLeuGlnAsnGluProThrValGluGluLeuValSerMetSerLeuLeuGly 110
Db 394 TTCAGCTAGAAAGTAGCCCACTGTGGAGGGCTTAATCGTGGCCATGTCACATGATGT 453
Qy 111 AlaThrIleValThrThrPheSerGlyProLeuSerAspSerIleGlyArgProMet 130
Db 454 GCAACCATCATCAACCAATCTCAGGCGGTATACAGCTGGATCGCGCGGCCCTATG 513
Qy 131 LeuIleLeuSerSerIleLeuLeuPhePheSerGlyLeuIleMetLeuTrpSerProAsn 150
Db 514 CTCATCTCTCTCAATCTCTACTTCTCAGCAGGCTCATCATGCTATGTTGGTCCGAT 573
Qy 151 ValTyrValLeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThr 170
Db 574 GTCTATGTGTACTGCTCGCAGCCTTATAGATGGATTCGGCATCGGCTTGGGTGTCAC 633
Qy 171 LeuValProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThr 190
Db 634 CTGTGACTTTGTACATCTCAGACAGCTCTCTTACAGATCAGGGGTTTGTCTGAATAC 693
Qy 191 LeuProGlnPheSerGlySerGlyMetPheLeuSerTyrCysMetValPheGlyMet 210
Db 694 CTGCCACAGTTCACTGAGTGGGTGAGGAGGATGTTCTTCTCTACTGCTGCTTGGATG 753
Qy 211 SerLeuSerProSerProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeu 230
Db 754 TCACCTGTGCGTCACTGAGTGGAGATTTATCGTTGGTGTCTCGCAATCCCTTCGTG 813
Qy 231 PhePhePheGlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerIys 250
Db 814 TTTTCTTCGGATTGCAATATTCTATCTGCCAGATCACCAGATGGCTGTGTCAGCAA 873
Qy 251 GlyArgMetAlaGluAlaIysValLeuGlnLysLeuArgGlyLysAspValSer 270
Db 874 GGGCGATGGCTCAGCAAGAGAGTATTACAAAATTTACGTGGGAGAGGATGTCAC 933
Qy 271 GlyGluLeuSerLeuLeuGluGlyLeuGluValGlyValAspThrSerIleGluGlu 290
Db 934 GGAGAAATGGCTCTCTTGTGTAGGTTTGGAGTTGGAGCTGACATCTCCATTGAGAG 993
Qy 291 TyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspIysGlu 310
Db 994 TACATCATCGGACCGGCTATAGAGCCAGCTGATGAGCATGTTGTTGATGGCGATAGGAC 1053
Qy 311 GlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGly 330
Db 1054 CAATAAACACTGACGGGCTGTAGAGGGCCAAATCATGATTTGCTCGACCTTCCAAAGGA 1113
Qy 331 ProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGln 350
Db 1114 CCCAGCATCTTGGAAAGTGTCTTCTCTATCATCTCGTCAATGGCAGCATGGTGAACCA 1173
Qy 351 SerValProLeuMetAspProIleValThrLeuPheGlySerValHisLeuAsnMetPro 370
Db 1174 AGCGTCCCACTTATGATCTTATAGTACGCTTTTCGGCAGTGTCCATGAGAACATGCT 1233
Qy 371 GlnAlaGlyIysSerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerVal 390
Db 1234 CATGCTGGAGGAAGCATCGGAGTACATTTGTTCCCACTTTGGTAGTATGTTAGTGTG 1293
Qy 391 ThrAspGlnHisAlaLysAsnGluInTrpAspGluGluAsnLeuHisArgAspGlu 410
Db 1294 ACAGATCAGACCCCAAGGTTGATCAATCGGATGAGGAACTTCAATAGGATGATGAG 1353
Qy 411 GluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeu 430
Db 1354 GAATACGCTGCTGATGCTGAGGTTGACTATGAAGCAATATGAGCAATATGAGGATGATG 1413
Qy 431 SerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySer 450

Db 1414 TCCCGACAGACCAGCGCGAGGAAGGAGGACATTGCCACCACTGCTCATCTGCGAAGT 1473
Qy 451 AlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyAspGlyValSerSerThr 470
Db 1474 GCTCTGAGTATGAGAAGAAGAGCTCTTGAAGAGGCTGCGAGGGGTGAGCAGCACT 1533
Qy 471 AspIleGlyGlyTyrTrpGlnLeuAlaTrpIleSerGluLysGluGlyGluAsnGly 490
Db 1534 GGCATTTGTGGGATGCGAGCTTGCATGGAATGGTTCAGAGGAGGAGGAGGATGGT 1593
Qy 491 ArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArg 510
Db 1594 AAGAAGGAGGAGGATTTAAAGGATCTACTTGCCACCAAGGAGGATGCCAGGATCAAGA 1653
Qy 511 ArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheVal 530
Db 1654 AGGGCTCAGTATTATTTCATCTCTGTTGGAGGAGCGCTCTCTGAAGGAGCGAATTCATA 1713
Qy 531 HisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArg 550
Db 1714 CATCTCGCGCTTGTGTAAAGCCACGACCTTTACTTCCAGGATATCATTTGACACCGT 1773
Qy 551 MetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLys 570
Db 1774 ATGTCGCGTCCAGCCATGATTCATCCATCAGAGGAGCTGCCAAAGGTTCAAGCTGAAA 1833
Qy 571 AspLeuPheGluProGlyValArgAlaLeuValGlyValGlyIleGlnIleLeu 590
Db 1834 GATTTGTTGACCTGGAGTGAGCGTGTGTTAGTCGGTGTGGAAATCAATTCCTT 1893
Qy 591 GlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluAla 610
Db 1894 CAACAGTTTGTGGAATAAATGGGTTCTTATTACACTCCACAAATACTTTGAGCAAGCG 1953
Qy 611 GlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIle 630
Db 1954 GGGTGGCAGTTCTCTTCCAAATCTTGGCTCAGTTCCAGTCCAGTCCATCTCCATCTGATC 2013
Qy 631 SerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAsp 650
Db 2014 AGTTCTCTGACCCCTACTGATGCTTCTTAGCATTTGGTTAGCCATGAGACTTTAGGAC 2073
Qy 651 LeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuVal 670
Db 2074 ATCTCTGGAAGAAGTTCTGCTTCTGGGCACAAATCCAGTCTTGATAGCATCTCAGTT 2133
Qy 671 IleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThr 690
Db 2134 GTCTTGGTGTGTTCCAAATGTTATCGACCTGGGTACAGTGGCCCGCCGCACTCTCCACA 2193
Qy 691 ValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIle 710
Db 2194 ATCAGCGTCACTCATCTACTTCTGCTGCTGCTCATGGGATTCGGTCCGATCCCAACAT 2253
Qy 711 LeuCysAlaGluIlePheProThrArgValArgGlyCysIleAlaIleCysAlaPhe 730
Db 2254 CTGTGTCGGAGATCTTCCCACTAGGTCGGCGGATCTGCAATGCCATCTGCGGCTG 2313
Qy 731 ThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIle 750
Db 2314 ACATTTCTGGATTGGTATATCATAGTCACTACAGCTCCCTGTGATGTGATGCGCATC 2373
Qy 751 GlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheVal 770
Db 2374 GGCTTACAGTGTCTTGGTATATAGCAGTGTGTTCGATTGCCCTTCGTTGCTGCTC 2433
Qy 771 PheLeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAla 790
Db 2434 TTCTCTCAAGTCCCGAGACGAAGGCGATGCCGCTCGAAGTCATCACCAGTCTCTCGCT 2493
Qy 791 ValGlyAlaLysGlnAlaAla 798
Db 2494 GTTGGTCAAAGCAAAATGCGAGCT 2517


```

RESULT 3
AY165599
LOCUS
DEFINITION
  AY165599 2665 bp mRNA linear PLN 07-JUN-2003
  Saccharum hybrid cultivar putative sugar transporter type 2a mRNA,
  complete cds.
ACCESSION
  AY165599
VERSION
  AY165599.1 GI:31505503
KEYWORDS
  Saccharum hybrid cultivar
  Saccharum hybrid cultivar
  Saccharum hybrid cultivar
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
  clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
  complex.
REFERENCE
  1 (bases 1 to 2665)
  Casu,R.E., Grof,C.P.L., Rae,A.L., McIntyre,C.L., Dimmock,C.M. and
  Manners,J.M.
  Identification of a novel sugar transporter homologue strongly
  expressed in maturing stem vascular tissues of sugarcane by
  expressed sequence tag and microarray analysis
  Plant Mol. Biol. 52, 371-386 (2003)
REFERENCE
  2 (bases 1 to 2665)
  Casu,R.E., Grof,C.P.L., Rae,A.L., McIntyre,C.L., Dimmock,C.M. and
  Manners,J.M.
  Direct Submission
  Submitted (17-OCT-2002) Division of Plant Industry, Commonwealth
  Scientific and Industrial Organisation, 120 Meiers Rd,
  Indooroopilly, QLD 4068, Australia
  Location/Qualifiers
    1..2665
      /organism="Saccharum hybrid cultivar"
      /mol_type="mRNA"
      /db_xref="taxon:129810"
    97..2334
      /codon_start=1
      /product="putative sugar transporter type 2a"
      /protein_id="AA037640.1"
      /db_xref="GI:31505504"
      /translations="MSGAAVAIVASIGNLQGDWNTASALLYIKFEFKLESPTV
      EGLVAMSLIGATIIITFGSPVSDWIGRLPMLILSSVLYFVSLLIMNSPNVYVLLLA
      RLINGVGVLAVLTVLIYISETAPPEIRGLNLFPOFSGSGMFIUSICMVGMSLLPS
      PDWRLMGVLADSLFFGLTIYPESPRWLVSIGRMAEKVKQLRGKDDVTGEM
      ALIVEGLVDGTFIEEYITGPAIEPADDDHADREIILFPEEGQSWIARPSKGS
      SMUGVSTESGRSGVMVONPLPTDPIVLTFSGSHENISLLAGSMERIVPSKFDMSIS
      ITDHPKTEOWDENLHRNEEVATDGAASDVNNVSHSLGRONTSAEGKIYAHGH
      RGSALGLRRSLSDGEAATSTGIGGWLAKWSEREGEDQKSGSKRVLVHOEG
      VAGSRGVSVPLEGGDASGGKFIHAALVSGSALYPRDIQEQMAGPATWHPSEAA
      AKVPSWRDLFEPGVRRALLVGIGIQLQFAGINGVLYIYTPQIMEQGVAVLISNLGL
      SSASILISSVTALLMLPSIGLMLMDVSGRRFLLSITPVLIASLIVLVSVNIE
      LGTVHVLSTISVITLCCFKMGFGPIINILCAEFPRVRGICITACALIFVWGD
      IVTYSLPVMLNAGLEGVGIYAVACAIAPFVYLKVPETKGMPLVITEFFAVGAQK
      AVAKA"
    97..156
      /note="putative; transmembrane-region site"
    115..774
      /note="pfam00083; Region: sugar transporter"
    238..297
      /note="putative; transmembrane-region site"
    313..372
      /note="putative; transmembrane-region site"
    406..462
      /note="putative; transmembrane-region site"
    511..570
      /note="putative; transmembrane-region site"
    589..654
      /note="putative; transmembrane-region site"
    658..1794
      /note="Region: large cytoplasmic loop"
    1624..2298
      /note="pfam00083; Region: sugar transporter"
    1672..1731
      /note="putative; transmembrane-region site"

```

```

misc_feature 1813..1869
  /note="putative; transmembrane-region site"
misc_feature 1897..1953
  /note="putative; transmembrane-region site"
misc_feature 1969..2022
  /note="putative; transmembrane-region site"
misc_feature 2089..2151
  /note="putative; transmembrane-region site"
misc_feature 2191..2247
  /note="putative; transmembrane-region site"

ORIGIN
Alignment Scores:
Pred. No.: 5,02e-201 Length: 2665
Score: 3155.00 Matches: 601
Percent Similarity: 89.56% Conservative: 68
Best Local Similarity: 80.46% Mismatches: 78
Query Match: 77.42% Indels: 0
DB: 8 Gaps: 0

US-10-051-909-32 (1-800) x AY165599 (1-2665)
QY 54 AsplysMetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuGln 73
  :::
Db 91 GAAACATGTCGGAGCTGCTCTTTGTTGCAATAGTAGCTTCGATCGGCAATCTCTTCAG 150
QY 74 GlyTTPAspAsnAlaThrIleAlaAlaValLeuTyrIleLysGluPheGlnLeu 93
  :::
Db 151 GGTGGGACATGCCACCTCTCAGCTCTATTATATATAAGAGGATTCAACTA 210
QY 94 GlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIle 113
  :::
Db 211 GAAAGTGAGCCTACTGTGGAGGGGCTGATTGTGGCCATGTCTCAATAATTGGACCACTATC 270
QY 114 ValThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeu 133
  :::
Db 271 ATTACACATTTCTTGGACCATATCAGACTGAGATGGCCGGCTCCCTATGCTCATCTC 330
QY 134 SerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrVal 153
  :::
Db 331 TCTTCAGTCTGTACTTTGTTAGCAGCCTCATTTATGCTATGCTCCCGAATGCTATGTA 390
QY 154 LeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValPro 173
  :::
Db 391 CTACTGCTGGACGCTCTTATAAATGATTCGGTTCGGCTGGCTGTCACCTCTTGTTCCCT 450
QY 174 LeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGln 193
  :::
Db 451 CTGTACATTTAGAAACAGCTCTCCACAGATAGAGGTTTGTGTAATACATGTCACAG 510
QY 194 PheSerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213
  :::
Db 511 TTCAGTGGGTTCAGGGGGGATGTTTTTGTCTACTGTCATGTTGCGGATGCTCATCTGCTG 570
QY 214 ProSerProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhePhe 233
  :::
Db 571 CCATCACCTGATTTGGAGGATATGTTGGGGTTCTCGGCTTACCTTCATTTGTTTCCTT 630
QY 234 GlyLeuThrIlePheTyrLeuProGluSerProArgTyrPheLeuValSerIleGlyArgMet 253
  :::
Db 631 GGTTCGACATATTTTATCTGCTCAATCACCAGATGGCTTGTGCAAAAAGGGCGAATG 690
QY 254 AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyLeuLeu 273
  :::
Db 691 GCAGAGGCAAAAGAGGTGTTGCAAAAGATTAGCTGGGAAGGATGACGTCAAGGCGAAATG 750
QY 274 SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIle 293
  :::
Db 751 GCTCTCTTGTGTAAGGGTTGGAGGTCGAGGGGATACCTTCATTGAAGAGTACATCAT 810
QY 294 GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThr 313
  :::
Db 811 GGTCTGCTATCGAACACGAGCTGATCACCGCTCGCTGATGTTGTTGATAGAACAGATAATA 870

```

Qy	314	LeuTyrGlyProGluGluGlyGlnSerThrPheAlaArgProSerLysGlyProIleMet	333	Qy	674	ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal	693			
Db	871	CTATTGGGCTGAAGAGGCGCAATCATGTAGTCCGGCTTCAAAGGGTCCAGATG	930	Db	1951	GTGCTAATGTATCGAGTTGGGTACAGTGGTCCATCTGTGCTCTCCACAATTAGTGTC	2010			
Qy	334	LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro	353	Qy	694	IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla	713			
Db	931	CTTGGAAGTGTGCTTCTATCGGATCTCGTCATCGGAGCATGTGTGAACCAAAATTTACCC	990	Db	2011	ATCACTACTTATCTGCTCTTCAAAATGGCTTTGGCCCATTCCTCCACATTTCTATGTGCA	2070			
Qy	354	LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly	373	Qy	714	GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTip	733			
Db	991	CTTACGGATCTATTGTACAGCTTTTGGAGTGTCTCATGAGAACATCTCTCTCGGTGGA	1050	Db	2071	GAGTTTTTTCACACCGAGTTCGTGGCATCTGCCATCTGTGCTTGATATTGG	2130			
Qy	374	GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln	393	Qy	734	IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla	753			
Db	1051	GGAAGTATGAGGAGAAATAGTGCCTTCTAAATTTGACAGTATGATCATGATCAGATCAG	1110	Db	2131	CTTGGAGACATTTATGTACATACAGTCTTCTGTGATGTGATGTGATGTGCTAGAA	2190			
Qy	394	HisAlaLysAsnGlnGlnTyrAspGluGluAsnLeuHisArgAspAspGluGluTyrAla	413	Qy	754	GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys	773			
Db	1111	CACCCCAAGACCGAGCAGTGGGATGAGGAGAACTCCATAGGGATATGAGGAGTATGCA	1170	Db	2191	GGCTGTTTGGCATATAGCAGTTGCTGTGCCATTCCTTTGTTGTTGTTGTTGTTTAAAG	2250			
Qy	414	SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln	433	Qy	774	ValProGluThrLysGlyMetProLeuGluValIleThrGluPheAlaValGlyAla	793			
Db	1171	ACTGATGCTGCTGCAAGTGACTATGAAACAAATGTACACAGCCCTCTGTGTCGCCAGCAG	1230	Db	2251	GTTCCGGAGACAAAGGCGCATGCCCTTGAGTGCATCTGAGTCTTCTTGGCAGTTGGTGCA	2310			
Qy	434	AlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGlySerAlaLeuSer	453	Qy	794	LysGlnAlaAlaAlaLysAla	800			
Db	1231	AATACAGTCCCGAAGGAGGAGCATTCACATCACGGGACCGTGGAAAGTCTTTGGGT	1290	Db	2311	AAGCAAGCGTGTGCAAGGCC	2331			
Qy	454	MetArgArgGlnThrLeuLeuGlyGluGlyGlyValSerSerThrAspIleGly	473	RESULT 4						
Db	1291	TTGAGAAGAGAGCCCTGTCGATGAGGTGTGAGGCGAGCAACAGCAGCTGCATTGT	1350	AK099716						
Qy	474	GlyGlyTyrGlnLeuAlaTyrLysTyrSerGluLysGluGlyGluAsnGlyArgLysGlu	493	AK099716.1	GI:32984925					
Db	1351	GGCGGATGGCAGTCCGCTGGAAATGTCAGAGCGAGAGGTGAGGACGGTAAAGAGAA	1410	VERSION	AK099716.1	GI:32984925				
Qy	494	GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgGlySer	513	KEYWORDS	FLI CDNA; CAP trapper.					
Db	1411	GGAGTTTCAAGAAATCTACTTACACCAAGAGGGAGTTGCTGGCTCAAGAGGGGATCT	1470	KEYWORDS	Oryza sativa (japonica cultivar-group)					
Qy	514	IleValSerLeuProGlyGlyAspValPheGluGlySerGluPheValHisAlaAla	533	SOURCE	Oryza sativa (japonica cultivar-group)					
Db	1471	GTTGTCTCACTCTCTGCTGAGGTGATGCTCTGAGGTGTTGAAGTTTATACATGCTGCT	1530	ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.					
Qy	534	AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp	553	REFERENCE	1					
Db	1531	GCTGTGTCAGCCAGTCAGCTCTTTACCGAGGGATATTACAGACAGCGCATGCCCGGT	1590	AUTHORS	The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team;					
Qy	554	AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTyrLysAspLeuPhe	573		Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yanagi, M., Suzuki, K., Li, C., Otsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Otono, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kuroaki, R., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, T., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., Riken, J., Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.					
Db	1591	CCAGCTACAAATGACCCATCAGAGCGCAGTGCCTGCAAGTTCCAAAGTGGAGACCTGTT	1650		Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice					
Qy	574	GluProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnIlePhe	593	JOURNAL	Science 301 (5631), 376-379 (2003)					
Db	1651	GAACTGGTGTGAGAGCTGCCCTGTTGGTGTGTTGTAATTCAGATCTCTCAACAGTTC	1710	MEDLINE	22752273					
Qy	594	AlaGlyIleAsnGlyValLeuTyrThrProGlnIleLeuGluAlaGlyValAla	613	PUBMED	12869764					
Db	1711	GGGGTAATAATGGGTTCTTACTACATCTCCGCAATTTATGAGCAAGCTGTGTGCA	1770	REFERENCE	2 (bases 1 to 2378)					
Qy	614	ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerLeu	633	AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, Y., Ikeda, R., Imamura, K., Inotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,					
Db	1771	GTTCTTATTCCAAATCTTGGTCTCAGTTCGGCATCAGCATCCTTAATTAGTTCTGTC	1830							
Qy	634	ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly	653							
Db	1831	ACTGCCCTACTTATGCTCCCTAGCATTTGTTAGCATGAGACTTATGAGCATCTCTGGA	1890							
Qy	654	ArgArgPheLeuLeuLeuGlyThrIleProIleLeuAlaSerLeuValIleLeuVal	673							
Db	1891	AGAAGGTTTCTGCTGTAAGCACAAATTCAGTCTTGTATAGCATCTTAAATTGCTCGTGT	1950							

Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niihara, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otono, Y., Ryu, R., Saichoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shiginawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : <http://cdna01.dna.affrc.go.jp/cDNA/>
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otono, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niihara, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carrinchi, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiroaka, T., Hori, J., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Satoh, K., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashiraki, Y.

Location/Qualifiers
1. 2378
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J01308619"

FEATURES

source
1. 2378
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J01308619"

ORIGIN

Alignment Scores:
Pred. No.: 3,17e-197 Length: 2378
Score: 3097.00 Matches: 596
Percent Similarity: 93.20% Conservative: 34
Best Local Similarity: 88.17% Mismatches: 46
Query Match: 76.00% Indels: 0
DB: 8 Gaps: 0

US-10-051-909-32 (1-800) x AK099716 (1-2378)

Qy 123 AspsrileGlyArgArgProMetLeuileuSerSerileLeuTyrPhePheSerGly 142
Db 1 GACTGGATCGCGCGCGCCCTATGCTCATCTCTCTCTCAATTTCTTACTTCTCAGCAGC 60
Qy 143 LeuileMetLeuTyrPsrProAsnValTyrValLeuLeuLeuAlaArgPheValAspGly 162
Db 61 CTCACTATGCTATGTTCCCGAATGTTCTATGTTGTTACTGCTCGCAGCCCTTATAGATGGA 120

Qy 163 PheGlyileGlyLeuAlaValThrLeuValProLeuTyrileSerGluileAlaProSer 182
Db 121 TTCGGCATCGCGCTTGGGTGTACACTTGTACCTTTGTACACTCTCAGAGACAGCTCTTCA 180
Qy 183 GluileArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeu 202
Db 181 CAGATCAGGGTTTGTCTGAATACACTGCCACAGTTTCACTGGTTCAGTGGTTCAGGAGGATGCTT 240
Qy 203 SerTyrCysMetValPheGlyMetSerLeuSerProSerProAspTyrArgIleMetLeu 222
Db 241 TCCTACTGCTATGCTTGGGATGCTACTGTTGCCGTCACTGCTGAGATTAATGCTT 300
Qy 223 GlyValLeuAlaileProSerLeuPhePheGlyLeuThrilePheTyrLeuProGlu 242
Db 301 GGTGTTCTCGCAATCCCTTCTGTTGTTTCTTCGGATGCAATATCTATCTCCAGAA 360
Qy 243 SerProArgTyrLeuValSerLysGlyArgMetAlaGluAlaLysValLeuGlnLys 262
Db 361 TCACCAAGATGGCTTGTTCAGCAAGGGCGGATGCTGAGGCAAGAGGATTTTCAAAA 420
Qy 263 LeuArgGlyLysAspAspValSerGlyLeuLeuSerLeuLeuGluVal 282
Db 421 TTACGTGGGAGAGGAGGATGCTTCAGGAGAAATGGCTCTTCTTGTGAGGTTTGGAGTT 480
Qy 283 GlyGlyAspThrSerileGluGluTyrileGlyProAlaThrGluAlaAlaAspAsp 302
Db 481 GGAGCTGACACTTCATTGAAGAGTATCATCTCGACCGGCTATAGAGCCAGCTGATGAG 540
Qy 303 LeuValThrAspGlyAspLysGluGlnileThrLeuTyrGlyProGluGluGlyLys 322
Db 541 CATGTTGTTGATGGGATTAAGGACCAATAACACTGTACGGCGCTGAGAGGGGCAATCA 600
Qy 323 TrpIleAlaArgProSerLysGlyProileMetLeuGlySerValLeuSerLeuAlaSer 342
Db 601 TGGATTGCTCGACCTTCCAAAGGGACCCAGCATCTTGGAGCTGTCTTCTTACATCT 660
Qy 343 ArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPhe 362
Db 661 CQTATGTCAGCATGGTGAACAGACGCTGACCTTATGATGCTTATAGTCAACGCTTTTC 720
Qy 363 GlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPhePro 382
Db 721 GGCAGTGTCCATGAGAACATGCTCATGCTGAGGAGAACATGCGGAGTACATTTGTTCCC 780
Qy 383 AsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTyrAspGlu 402
Db 781 AACTTTGGTAGTATGTTTAGTGTGACACATCAGCACCCCAAGGTTGATCAATGGGATGAG 840
Qy 403 GluAsnLeuHisArgAspAspGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGlu 422
Db 841 GAGAACCTTCATAGGGATGATGAGGAATACCGCTCTGATGCTGAGGAGTACATGAA 900
Qy 423 AspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIle 442
Db 901 GACATGTCACAGCCCGTTGCTGTCCGACAGACCCAGCGCAGAGGAGGAGACATT 960
Qy 443 ValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGlu 462
Db 961 GCACACCATGCTCATCGTGGAGTGTCTCTGATGATGAGAGGAGGAGGCTCTTGAAGAG 1020
Qy 463 GlyGlyAspGlyValSerSerThrAspIleGlyGlyTyrGlnLeuAlaTyrLysTyr 482
Db 1021 GGTGCGAGGCGGTGAGCAGCATGCGCATTTGGTGGGATGGCAGCTTGCATGGAATGG 1080
Qy 483 SerGluLysGluGlyLysGlyArgLysGluGlyGlyPheLysArgValTyrLeuHis 502
Db 1081 TCAGAGCGAGAGGAGGAGGATGTTAAGAGGAGGAGGATTTAAAGAGGATCTACTTGCAC 1140
Qy 503 GlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAsp 522
Db 1141 CAAGAGGAAGTCCCGAGGATCAAGAAAGGGGCTCAGTTATTTTTCATCTTCTGTTGGAGGGAC 1200

Qy 523 ValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPhe 542
 Db 1201 GCTCCTGAGGACGAGTAATCATACATGCTGCGCTTTGGTAAGCAACGAGCACTTTAC 1260
 Qy 543 SerlySGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluVal 562
 Db 1261 TCCAAGGATATCATTTGAACGCGTATGTCGGTCCAGCCATGATTCATACATCAGAGGCA 1320
 Qy 563 AlaAlaGlySerArgTgTlyAspLeuPheGluProGlyValArgAlaLeu 582
 Db 1321 GCTGCCAAGGTTCAAGCTGGAAGAATTGTTGAACTGGATGAGGCGTGTTGTTA 1380
 Qy 583 ValGlyValGlyLeuGlnLeuGlnGlnPheAlaGlyLeuValLeuTyrTyr 602
 Db 1381 GTCGGTGTGGAATTCAAATTCCTTCAACAGTTGCTGGATTAATGGGTCTCTATTAC 1440
 Qy 603 ThrProGlnLeuLeuGlnAlaGlyValAlaValLeuLeuSerlySGlyLeuSer 622
 Db 1441 ACTCCAAATATCTGAGCAAGCGGGTGGCAGTTCTCTGACCACTTCTGATGTTCTAGCAT 1500
 Qy 623 SerAlaSerAlaSerLeuLeuSerSerLeuThrThrLeuLeuMetLeuProCysIle 642
 Db 1501 TCAGCATCAGTCCATCTTGATCAGTTCTGACCACTTCTGATGTTCTAGCAT 1560
 Qy 643 GlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIle 662
 Db 1561 GGTTTAGCATGAGACTTATGACATCTCTGGAAGAAGGTTCTGCTTCTGGGCAAT 1620
 Qy 663 ProIleLeuLeuAlaSerLeuValLeuValValSerAsnLeuLeuLeuGlyThr 682
 Db 1621 CCACTTGTATAGATCTCTAGTGTGCTGTGTGTGTCCTGCAATGATTCACCTGGGTACA 1680
 Qy 683 LeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMet 702
 Db 1681 GTGGCCCAAGCCGCACTCTCCCAATCAGCGTCATCATCTACTTCTGCTTCTGTCATG 1740
 Qy 703 GlyPheGlyProIleProAlaLeuLeuCysAlaGluLeuPheProThrArgValArgGly 722
 Db 1741 GGAATGGTCCGATCCCAACATCTGTGTGGAGATCTTCCCACTAGGTCGCGCGC 1800
 Qy 723 LeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleValThrTyrSer 742
 Db 1801 ATCTGATTCGATCTGCGGCTGACATCTCGATTGGTGGTATCATATCATCTACCTACAGC 1860
 Qy 743 LeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValVal 762
 Db 1861 CTCCTGTGATGTCGATGTCATGCGCTAGCAGGTGCTTTGGTATATACGACGTTGTT 1920
 Qy 763 CysLeuIleSerPheValPheValPheLeuLeuValProGluThrlySGlyMetProLeu 782
 Db 1921 TGTCTGATGCTTGTGTGCTTCTTCTCAAGTCCCGAGACGAGGCGATGCGGCTC 1980
 Qy 783 GluValIleThrGluPhePheAlaValAlaGlyAlaLysGlnAlaAla 798
 Db 1981 GAAGTCATCACCAGTCTTCTGCTGTGTGTGTGTCGAAAGCAATGACGCT 2028
 RESULT 5
 AP005756 136267 bp DNA linear PLN 29-JUN-2004
 LOCUS
 DEFINITION
 Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
 BAC clone:OSJNB0035N08;
 ACCESSION
 AP005756
 VERSION
 AP005756.3 GI:49388934
 KEYWORDS
 Oryza sativa (japonica cultivar-group)
 SOURCE
 ORGANISM
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 REFERENCE
 1
 AUTHORS
 Sasaki, T., Matsumoto, T. and Katayose, Y.
 TITLE
 Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC
 clone:OSJNB0035N08

JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Published Only in Database (2002)
 2 (bases 1 to 136267)
 Sasaki, T., Matsumoto, T. and Katayose, Y.
 Direct Submission
 Submitted (18-SEP-2002) Takuji Sasaki, National Institute of
 Agrobiological Sciences, Rice Genome Research Program, Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
 Tel: 81-298-38-7441, Fax: 81-298-38-7468)
 On Jun 28, 2004 this sequence version replaced gi:42627749.
 Genes were predicted from the integrated results of the following:
 GENSCAN (http://ccr-081.mit.edu/GENSCAN.html), FGENESH
 (http://www.softberry.com/), GeneMark.hmm
 (http://opal.biology.gatech.edu/GeneMark/), GlimmerM
 (http://www.tigr.org/tdb/glimmer/glmr.form.html), RiceHM
 (http://rgp.dna.affrc.go.jp/RiceHM/), SplicePredictor
 (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4
 (http://globin.cse.psu.edu/html/docs/sim4.html), gap2
 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The
 genomic sequence was searched against NCBI NonRedundant Protein
 database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
 database at RGP or DDBJ. Protein homologies of the coding
 regions were searched against NCBI NonRedundant Protein database
 with BLASTP. ESTs represent the identified cDNA sequences using
 BLASTN with the corresponding DDBJ accession no. and RGP clone ID.
 Full-length cDNAs represent the identified cDNA sequences using
 BLASTN with the corresponding DDBJ accession no.
 A gene with identity or significant homology to a protein is
 classified based on the protein name to indicate the homology level
 such as same name, 'putative' and '-like protein'. A gene without
 significant homology to any protein but with full-length cDNA or
 EST homology (covering almost the entire length of partial
 sequence) is classified as an 'unknown' protein. A gene predicted
 by two or more gene prediction programs is classified as a
 'hypothetical' protein according to IRGSP standard. A gene
 predicted by a single gene prediction program is also classified as
 a probable 'hypothetical' protein and is included as a
 miscellaneous feature of the sequence.
 The orientation of the sequence is from M13rev to -21M13 of the BAC
 clone. This sequence of OSJNB0035N08 clone has an overlap with
 P0620H05 (DDBJ: AP005394) clone at 5' end and with QJ171.D06
 (DDBJ: AP004857) clone at 3' end. Detailed information on overlap
 and assembly quality together with annotation of this entry is
 available at http://rgp.dna.affrc.go.jp/GenomeSeq.html.
 FEATURES
 location/Qualifiers
 1. 136267
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="2"
 /clone="OSJNB0035N08"
 /complement(join(157..256,518..608,645..687,1665..1820))
 /gene="OSJNB0035N08.1"
 /complement(join(157..256,518..608,645..687,1665..1820))
 /gene="OSJNB0035N08.1"
 /note="hypothetical ORF
 predicted by GENSCAN
 this category is not included in IRGSP standard"
 /complement(6895..7272)
 /gene="OSJNB0035N08.2"
 /complement(6895..7272)
 /gene="OSJNB0035N08.2"
 /note="hypothetical ORF
 predicted by GlimmerM
 this category is not included in IRGSP standard"
 /complement(11805..13042)
 /gene="OSJNB0035N08.3"
 /complement(join(11805..12148,12855..13042))
 /gene="OSJNB0035N08.3"
 /note="supported by full-length cDNA(s): AK111070"
 /complement(11805..13042)
 /gene="OSJNB0035N08.3"
 gene
 misc_feature
 gene
 misc_feature
 gene
 mRNA
 misc_feature

[illegible]

Qy	81	-----	81	-----	81
Db	48860	GTCTCGTGTAGATTGGTTATTGGTTGAGAGTTTGTGAGTACTAGGAAATATGAACAGC	48919		
Qy	81	-----	81	-----	81
Db	48920	TGTCAAATTTTTCCTTAATTAATCAGTGTGGTGATCCTTCTAAATTTTTCAGATTAATAGA	48979		
Qy	81	-----	81	-----	81
Db	49039	CTGCTTTAGATCAGATGTTTGTGTGAGTGTCTAATTAATGTCCTTAGTAACATGTCGTGTG	49039		
Qy	81	-----	81	-----	81
Db	49040	TTGTTGTTGAGTCTPAAGAGCAATTTGGCACAAATGTTTATCAGCAAACTGGATCGCATG	49099		
Qy	81	-----	81	-----	81
Db	49100	ATATTAAATAGTCTCTTAATTTTTCCTAGATCTGCTACTTATTGAAGAAAGAGATGTT	49159		
Qy	81	-----	81	-----	81
Db	49160	CTAATCATAATTTGTTAGCTATTGTGTGTCTGTGAATCATTACTGTTTGGCAATCTTTTG	49219		
Qy	81	-----	81	-----	81
Db	49220	TTCTGAACACCTGTAAAGATTGTTCTATCCAACAAGAAAGTGTAATGTTTAAACAGGCAC	49279		
Qy	82	-----	AlaalaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn	95	
Db	49280	AATTATTAACCTTCTTCAGGTGCTGTACTATACATAAAGACGAATTCAGAGCTAGAAAGT	49339		
Qy	96	GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyValThrIleValThr	115		
Db	49340	GAGCCCACTGTGAGGGGTAACTGGGCCATGTCACTGATTGGTGCAACCATCATCA	49399		
Qy	116	ThrPheSerGlyProLeuSerAspSerIleGlyA-gArgProMetLeuIleLeuSerSer	135		
Db	49400	ACATTCTCAGGGCCGGTATCAGACTGGATCGCGCGGCCCTATGCTCATCTCTCTTCA	49459		
Qy	136	IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu	155		
Db	49460	ATTCTCTACTTCTCAGCAGCCTCATCATGCTATGTCCTCCGAATGTCATGTGTACTG	49519		
Qy	156	LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr	175		
Db	49520	CTCGCAGCCCTATAGATGGATTCGGCATCGCGCTGGCTGTGCACACTTGTACCTTGTAC	49579		
Qy	176	IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer	195		
Db	49580	ATCTCAGACAGACTCCTTCAGAGATCAGGGGTTTGCTGAATACATACCTGCCACAGTTCAGT	49639		
Qy	196	GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer	215		
Db	49640	GGGTCAAGAGGGATGTTCTTGCTACTGCTGCTGGTGGGATGTCATGTTGCCGTCA	49699		
Qy	216	ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu	235		
Db	49700	CCTGACTGGAGAATTATGTTGGTGTCTCGCAATCCCTTCGTTGTTTCTTCGGATTG	49759		
Qy	236	ThrIlePheTyrLeuProGlnSerProArgTrpLeuValSerLysGlyArgMetAlaGlu	255		
Db	49760	ACAAATTTCTATCGCGAAATCACAGATGGCTGTGACGAAAGGGCGGATGGCTGAG	49819		
Qy	256	AlaLysLysValLeuGlnLysLeuArgGlyLysAspValSer-	270		
Db	49820	GCAAGAAAGTATTACAAAAATTACGTGGGAGAGGATGTCCTCAGGATGTTTGTGCTGA	49879		
Qy	270	-----	270	-----	270
Db	49880	AAATTTTGATGTTCAATTCACCCATATTTCCAGCTGACAGGTTTAAAAACAATTTCTCCC	49939		

Qy	270	-----	270	-----	270
Db	49940	AGATTCTAAACTCATGGTTTAATACTAGAAAACATGTGGCATTTGGTAACCTGTTAAATTTTC	49999		
Qy	270	-----	270	-----	270
Db	50000	TTGACAGCACTTTAACAATATGCAGAAATAGTGTATTTTGGAAAAAAGGCTAAAAAGAA	50059		
Qy	270	-----	270	-----	270
Db	50060	TCAAAATCTATTAAAGGCCTTCTATCGGTGCTCTTTTGGAAAAAGCTTATGAATCTTCAAG	50119		
Qy	270	-----	270	-----	270
Db	50120	TAAAAATTACTAAGTACACATAACTACATAATTCATTTTCCATAGAAGGCAGACATGG	50179		
Qy	270	-----	270	-----	270
Db	50180	AATGTACAATCGTTATCTCACAGGTAGAGCCAAATGAGGTGTCCCAACTTTTGTCTGA	50239		
Qy	270	-----	270	-----	270
Db	50240	GTCTTAATCTGGAGAAAAATAATCAGGTGATCTAGATAGACATCGCCCTCTTTTTTTTTT	50299		
Qy	270	-----	270	-----	270
Db	50300	TTTTGGTAAGGTAACCTCAGAAATTTGGGGGTGGACCTCGTGTATAATGTGATGTG	50359		
Qy	270	-----	270	-----	270
Db	50360	GGCGGACCCAGGTCCTTATCTACTCTAGTTTAACTAAATGAACATAATTTCTCTTTG	50419		
Qy	270	-----	270	-----	270
Db	50420	AGATTTTAAAGTGAATTAAGTCCATCTAGTTGNACCTAGAGTGGGTCAAGGAGCCCCAC	50479		
Qy	270	-----	270	-----	270
Db	50480	AGCGGCATCGCTAAACAACAGTGAATTGTTCAAAAAGTTAGTAATGATTTATCCTTAAGT	50539		
Qy	270	-----	270	-----	270
Db	50540	TCATGCTTTCCTCAAACTAATTACATAATACATGTGTCAACATCAAAATAGGCATATGCT	50599		
Qy	270	-----	270	-----	270
Db	50600	GTTAGCAAGTGGTTTCTCTATTTTCTCCTCTACAACAGACTAAATGGACAGTATTCGGTT	50659		
Qy	270	-----	270	-----	270
Db	50660	GGCTGAGACATAATATTGCTGTATTTTAAGTTCAATCGAACTTCGGTGTAAAGCTGGTTG	50719		
Qy	270	-----	270	-----	270
Db	50720	TCCCCAGCTGCATATGCTAGCTTTTCTATTAAAGCTGGGCCCTTATATTTCTGTATTTG	50779		
Qy	270	-----	270	-----	270
Db	50780	TTTCTGGCACACCATTCATATTATATTGCTTTATTCATAGTTACTACTTATTTATG	50839		
Qy	270	-----	270	-----	270
Db	50840	CCAAATGTTTATACCAATGTCBAAATAAATCATGTAACTCTACGTGAAAAATAAAGTTGC	50899		
Qy	271	-----	281		
Db	50900	ATAATCATTCAAATTAATCAAAATGCGCAGAGAAATGGCTCTTCTGTTGAAGGTTTGA	50959		
Qy	281	uValGlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaIleAs	301		
Db	50960	GGTGTGAGCTGACATCTTCATTGAAGATACATCATCGACCGGCTATAGAGCAGCTGA	51019		
Qy	301	pAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlyG	321		

Db 51020 TGAGCATGTTGTTGATGGGATAAGACCAATAACACACTGTACGGGCTTGAAGAGGCCA 51079
Qy 321 nSerTrpIleAlaAArgProSerIysGlyProIleMetLeuGlySerValLeuSerIleuAl 341
Db 51080 ATCATGGATTGCTCGACCTTCCAAAGGACCCAGCATCTTGGAAAGTGTCTTCTTAC 51139
Qy 341 aSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLe 361
Db 51140 ATCTCGTCATGGAGCATGGTGACAGAGCGTGCCACTTATGGATCTTATAGTCAAGCT 51199
Qy 361 uPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPh 381
Db 51200 TTTCCGCGAGTCCATGAGAACATGCTCATGTGAGGAAGCATCGGAGTACATTGTT 51259
Qy 381 eProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAs 401
Db 51260 TCCCAACTTTGGTAGTATGTTAGTGTGACAGATACAGACCCCAAGGTTGATCAATGGAA 51319
Qy 401 pGluGluAsnLeuHisArgAspAspGluGluTrpAlaSerAspGlyAlaGlyGlyAspTy 421
Db 51320 TGAGGAGAACCTTCATAGGAGTATGAGGAATACGGCTCTGATGTGCTGGAGTGACTA 51379
Qy 421 rGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAs 441
Db 51380 TGAAGACAATGTCCACAGCCGCTGTGCTGCCACAGACACACAGCGCAGAGGGAAGGA 51439
Qy 441 pIleValHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnAlaThrGlyAlaGluGlyLysAs 461
Db 51440 CATTCACACCATCTCATCTGTGAGTGTCTGATGATGAGAAAGAACGCTCTTGA 51499
Qy 461 yGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpL 481
Db 51500 AGAGGGTGGCGGCGGTGAGCAGCATCTGCATGTGTGGGGATGCGAGCTTGCATGGAA 51559
Qy 481 sTrpSerGluLysGluGlyGlnAsnGlyArgLysGluGlyGlyPheLysArgValTrpLe 501
Db 51560 ATGCTCAGACGAGAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 51619
Qy 501 uHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyG 521
Db 51620 GCACCAAGAGAGAGTGCAGGATCAGAGAGGCGTCACTATTCTTCTCTGCTGGAGG 51679
Qy 521 YAspValPheGluGlySerGluPheValHisAlaAlaAlaLeuValSerGlnSerAlaLe 541
Db 51680 GAGCGCTCCCTGAAGGCGAGCAATTCATACATGCTGCCGCTTGTGTAAGCCACACGACT 51739
Qy 541 uPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGl 561
Db 51740 TTACTCCAGGATATCATTTGACAGCGTAATGCCGTCAGCCATGATTCATCCATCAGA 51799
Qy 561 uValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLe 581
Db 51800 GGCAGCTGCCAAAGGTTCAAGCTGGAAAGATTTGTTTGAACCTGGAGTGAGGCGTGGTT 51859
Qy 581 uLeuValGlyValGlyIleGlnIleLeuGln----- 591
Db 51860 GTTAGTCGGTGTGGAAATTCATATCCCTTCACAGGTAGATGATACCAATTTCAATAATT 51919
Qy 591 ----- 591
Db 51920 CTTTTTTGTACTGTCCAAACATTTGACGCTGTTAAGATTTGTGACTATTAAAGAATTCAC 51979
Qy 592 -----GlnPheAlaGlyIleAsnGlyValLeuTrpTr 603
Db 51980 AATTCCCTTTTANGTCTTTCTGCTGAGTTGCTGGAAATAATGGGGTCTCTATTATACA 52039
Qy 603 hrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerS 623
Db 52040 CTCCACAAATACTTGAAGAGCGGGTGGCAGTTCTCTCTCCAACTTGGCCTCAGTT 52099
Qy 623 erAlaSerAlaSerIleLeuLysSerSerLeuThrLeuLeuMetLeuProCysIleG 643

Db 52100 CAGCATCAGCTTCATCTTGATCAGTTCTCTGACCCACCTACTGATGCTTCCTAGCATTTG 52159
Qy 643 lyPheAlaMetLeuLeuMetAspLeuSerGlyArg----- 654
Db 52160 GTTAGCCATGAGACTTATGACATCTCTGGAAGAGGTACTAACTTGTTCCTTGGTTT 52219
Qy 654 ----- 654
Db 52220 GCATCATCTTAGTACTACCAATATATCAAAACAAGAAATGCACCAAGTTAGTATTAAAT 52279
Qy 655 -----ArgPheLeuLeuLeuGlyThrIleProIleLeuIleA 667
Db 52280 ACATCATACATTGTTCTCTTTTAGGTTTCTGCTCTTGGGACAAATTCAGCTTTGATAG 52339
Qy 667 laSerLeuValIleLeuValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaL 687
Db 52340 CATCTAGTTGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 52399
Qy 687 euLeuSerThrValSerValIleValTyPheCysCysPheValMetGlyPheGlyProI 707
Db 52400 CACTCTCCACAATCAGCGTCTCATCTTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTG 52459
Qy 707 leProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaI 727
Db 52460 TCCCAACATCTGTGTGCGGAGATCTTCCCACTAGGTCGCGGCATCTGCATTGCCA 52519
Qy 727 leCysAlaPheThrPheTrpIleGlyAspIleIleValThrTySerIleuProValMetL 747
Db 52520 TCTGCGCCCTGACATCTCGGATTCGATATCATAGTACCATCAGCTCCCTGCTGATGC 52579
Qy 747 euAsnAlaIleGlyLeuAlaGlyValPheSerIleTyAlaValAlaValCysLeuIleSerP 767
Db 52580 TGAATGCCATGGCTACGAGTGTCTTTGGTATATATAGCAGTTGTTGCTCGATTGCT 52639
Qy 767 heValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThrG 787
Db 52640 TCGTGTCTGCTTCTCCTCAAGTCCCGAGACGAGGATCCCGTCGAAGTCATCACCG 52699
Qy 787 luPhePheAlaValGlyValLysGlnAlaAlaA 798
Db 52700 AGTTCTTCGCTGTTGTCGAAAGCAATGCAAGCT 52734

RESULT 6
LOCUS AK102640
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone J033100A10, full insert sequence.
ACCESSION AK102640
VERSION AK102640.1 GI:32987849
KEYWORDS FLI CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
1
REFERENCE The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team:., Kikuchi, S. Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yasaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yanagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shiishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group:., Otonari, Y., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayaishizaki, Y.

TITLE

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 (5631), 376-379 (2003)
22752273
12869764
2 (bases 1 to 2800)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramatsu, K., Hiraoka, T.,
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Iotani, K.,
Iotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, T., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, K., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, K.,
Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,
Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
Osato, N., Ota, Y., Otonari, Y., Ryū, R., Saitoh, H., Sakai, C., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
Yanada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
Yoshimura, A.

Direct Submission

Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kanondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,
Tel: 81-29-838-7007, Fax: 81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica
rice.

COMMENT

URL : <http://cdna01.dna.affrc.go.jp/cdna/>
NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Naniki, T.,
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and
Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otonari, Y., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryū, R., Sugano, S.,
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramatsu, K.,
Hiraoka, T., Hori, F., Iida, J., Imamura, K., Iotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,
Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers

1. 2800
/organism="Oryza sativa (japonica cultivar-group)"
/mol type="mRNA"
/cultivar="Nipponbare"
/db xref="taxon:39947"
/clone="J033100A10"

FEATURES

source

Percent Similarity: 82.32% Conservative: 78
Best Local Similarity: 72.26% Mismatches: 121
Query Match: 69.34% Indels: 16
DB: 8 Gaps: 7
US-10-051-909-32 (1-800) x AK102640 (1-2800)
QY 30 AlaLeuProGlyProLeuProProAlaSerCys-----SerSerGI 43
Db 117 GCCTCTCCAGATCCAGCGCCCTCTCTCTTTAGGGATCCGAATTCGGGTGACG 176
QY 43 nGluProValThrSerAspAspIleLeuGluAspLysMetSerGlyAlaValLeuAla 63
Db 177 AGAGACTTGGTGGTAAGAT-----TCGCCGCCATGGCGGCCCTGCTGGTGGC 227
QY 63 aIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIleAlaAla 83
Db 228 CATCGCGCCCTCCATCGGCACTTGTCTGCGGGCTGGATATGCAACCATTCAGGTGC 287
QY 83 aValLeuTrpIleLysLysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeu 103
Db 288 GGTACTGTATCATCAAGAAGGAATTCACATTTCGAGCGAGCCCTTATCGAAGCCGTGAT 347
QY 103 eValSerMetSerLeuIleGlyAlaThrIleValThrPheSerGlyProLeuSerAs 123
Db 348 CGTGCCCATGTGCTCATTTGGGGCGACGATCATCAGCGCTTCTCTGAGCAGTGGCTGA 407
QY 123 pSerIleGlyArgProMetLeuIleLeuSerSerIleLeuTrpPhePheSerGlyLe 143
Db 408 TTTCTTTGGTAGGGGGCCCATGCTGATCGCGTGGCTGCTCTTCTTCTTCTTCTTCTTCT 467
QY 143 uIleMetLeuTrpSerProAsnValTrpValLeuLeuAlaArgPheValAspGlyPh 163
Db 468 AGTATGCTTTGGGGCCCAATGATGTGTCTTGGCGAGGCTCATTCAGCGGTT 527
QY 163 eGlyIleGlyLeuAlaValThrLeuValProLeuTrpIleSerGlyIleAlaProSerGI 183
Db 528 CGGGATCGGTTGGCTGTACGCTTGTACCATTTGTATCATCTCTGAGACTCCCCGACGGA 587
QY 183 uIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMetPheLeuSe 203
Db 588 CATCAGAGGACTGCTAAACACGTCGCCAGTTCAGTGGGTCTGGAGGGATGTTCCTTC 647
QY 203 rTrpCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMetLeuGI 223
Db 648 ATACTGATGGTATTGGCATGTCCTCATGCCACAGCCAGATTGGAGGATCATGCTGG 707
QY 223 yValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTrpLeuProGluSe 243
Db 708 CGTTCTATCAATCAATCACTTATATCTTGTGATTGACCATCTTTTACTTACTCTGAATC 767
QY 243 rProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLe 263
Db 768 GCCGAGTGGCTCGTGAGCAAGAGAGATGGCTGAGGCAACCGTGTGTTCAGGCCT 827
QY 263 uArgGlyLysAspValSerGlyGluLeuSerLeuLeuLeuGluGlyLeuValGI 283
Db 828 GCGTGGAGAGAGATGTTTTCAGGAGAAATGGCCCTTCTCGTTGAAGGTCTGGGGGTGG 887
QY 283 yGlyAspThrSerIleGluGlyTrpIleIleGlyProAlaThrGluAlaAlaAspLe 303
Db 888 GAAGACACAAAAATGAGGAATACATATGACCTGTATGATGAGCTGTGCTGATGAAGG 947
QY 303 uValThrAspGlyAspLysGluGlnIleThrLeuTrpGlyProGluGluGlyGlnSerTr 323
Db 948 GCTGGCTCCA--GATCCAGAGAGATCAAACTGTATGGTCTCTCAAGAGGCTTATCGTG 1004
QY 323 pIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerAr 343
Db 1005 GGTGGCCCGCTGTTACACGGCAAGATGCACATTCGAAGTGCATTAGGCTCATCTCTCG 1064
QY 343 gHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGI 363

ORIGIN

Alignment Scores:
Pred. No.: 2800
Score: 5.39e-179 Length: 2800
Matches: 2825.50

Db 1065 TCATGCTAGTGTGCTCAGTCAGGTAAGCCCTTGTGGATCTCTGTGTCAACCTTTTGG 1124
Qy 363 yserValHisGluAsnMetProGlnAlaGlySerMetArgSerThrLeuPheProAs 383
Db 1125 AAGTGTCCTATGAGAGATCCCTGAGATATGGAAGCATGCGGAGACATTTTCTTAA 1184
Qy 383 nPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluG 403
Db 1185 CTTTGGCAGCATGTTTGTGTGGCGAAGCAGCAGCAAGCTAAAGGTGATTGGATGCTGA 1244
Qy 403 uAsnLeuHisArgAspGluGluTrpAlaSerAspGlyAlaGlyAspTrpGluAs 423
Db 1245 GAGT---CAACGGGAGGCTGAAGATTATGATCAGACCATGTTGGGATGATCAATGAAGA 1301
Qy 423 pAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspLe 442
Db 1302 TAGCCTCCAAAGCCCACTTATTTCTCTCAAGCACAAGCTGGGAAGGAAGATCGC 1361
Qy 443 -ValHisHisGlyHisArgGlySerAlaLeuSerMetArgA:GlnThrLeuLeuGlyG 462
Db 1362 TGCACCTCATGGCAGTATAATGGTCTCTGGGA-----AGAAGTAGTAGTCTCATGCA 1415
Qy 462 uGlyGlyAspGlyValSerSerThrAspGlyGlyGlyTrpGlnLeuAlaTrpLeuTr 482
Db 1416 GGGCGGGAGGCGAGTAAGCAGCATGGCATTTGGTGGGATGCGATTTGGTTGGAATG 1475
Qy 482 pSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTrpLeuHi 502
Db 1476 GACTGAGAGAGAGGTGCAGATGGCGAAAGAGGTGGCTTCCACGATCTACTTGCA 1535
Qy 502 sGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyLys 522
Db 1536 TGAAGAGGTGTGACAGGTGATCGCAGGGGCTCTATCTGTCTATGTCCT---GGAGTGA 1592
Qy 522 pValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPh 542
Db 1593 TGTTCCTCTGGTGGTGGTTCCTCAGCAGCTGCTCTGTGAGCAACCTGCTCTTTA 1652
Qy 542 eSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluVa 562
Db 1653 CTCTAAGGAATTGATGAAGCAAGCTTGTGGCCCTGCTGTATGTCATCCATCTCAGGC 1712
Qy 562 LAAlaAlaGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLe 582
Db 1713 AGTTGCTAAAGGTCCTAAAGTGGCAGCTTATTCGAACCTGGAGTGAAGATGCTCTGT 1772
Qy 582 uValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTrpY 602
Db 1773 TGTTCGATAGGATACAAATCCTGCAACAGATTTGCTGGCATTAAAGGATCTGTACTA 1832
Qy 602 rThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSe 622
Db 1833 CACTCCACAAATCTTGAGCAAGCTGGTGTGTGTCTCTTCTTTCGAAACATTGGACTTAG 1892
Qy 622 rSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysI 642
Db 1893 CTCTCATCTGCATCTATTCTTATAGGAGCTGACAACTTCTGACGCTCCCAAGCAT 1952
Qy 642 eGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrI 662
Db 1953 TGGTATTGCTATGAGGCTCATGATATGCTGGAAGAAGTTTCTTCTCTTCTGCAACAAT 2012
Qy 662 eProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyTh 682
Db 2013 CCTATCTCTGATAGTACACTAGTATCTTGTATCTGGTCAATATCTGTGATGTGGGAC 2072
Qy 682 rLeuAlaHisAlaLeuLeuSerThrValSerValIleValTrpPheCysCysPheValMe 702
Db 2073 CATGGTTTCATGCTCATCTGCTCCACAGTCAGTGTCTACTCTCTCTCTTCTGCTAT 2132
Qy 702 tGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgG 722
Db 2133 GGGGTTGGGCGCTATTCCAAACATTCTCTGTGAGAGATTTTCCCGACCAACCGTTCTGG 2192

Qy 722 yLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTrpSe 742
Db 2193 CATCTGATAGCCATCTGTGCCCTAAACATTCTGGATCGGTGATATCATTTGTGACATAC 2252
Qy 742 rLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTrpAlaValVa 762
Db 2253 CCTCCCGTGATGTCTCAACCCCATTTGACTCGCTGGAGTGTGTTGGAATCTACCGAGTGT 2312
Qy 762 lCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMetProLe 782
Db 2313 CTGCTACTGCTTCTCTGTTTCTTCATGAAGTCCCGGAGACAAAGGCGATGCTCT 2372
Qy 782 uGluValIleThrGluPheAlaValGlyAlaLysGlnAla 796
Db 2373 TGAAGTCATCACCAGTCTCTCTGTGCGGACCAAGCAGGCC 2415
RESULT 7
HVU534445 2614 bp mRNA linear PLN 14-FEB-2003
LOCUS Hordeum vulgare mRNA for hexose transporter (stpl gene).
DEFINITION
ACCESSION AJ534445
VERSION AJ534445.1 GI:26986185
KEYWORDS hexose transporter; stpl gene.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poaceae; Triticeae; Hordeum.
REFERENCE 1
AUTHORS Weschke, W., Panitz, R., Gubatz, S., Wang, Q., Radchuk, R., Weber, H. and
Wobus, U.
TITLE The role of invertases and hexose transporters in controlling sugar
ratios in maternal and filial tissues of barley caryopses during
early development
JOURNAL Plant J. 33 (2), 395-411 (2003)
MEDLINE 22424051
PubMed 12535352
REFERENCE 2 (bases 1 to 2614)
AUTHORS Radchuk, R.
TITLE Direct Submission
Submitted (09-DEC-2002) Radchuk R., Molecular Genetics - Gene
Expression, Plant Genetics and Crop Plant Research, Corrensstr. 3,
Gatersleben, D-06466, GERMANY
FEATURES
source
1..2614
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/sub_species="vulgare"
/db_xref="taxon:112509"
/tissue_type="developing caryopsis"
/clone_lib="lambda ZAP2"
/dev_stages="0-15 days after flowering"
1..2614
/gene="stpl"
83..2314
/gene="stpl"
/codon_start=1
/product="hexose transporter"
/protein_id="CAD8958.1"
/db_xref="GI:26986186"
/db_xref="GOA:Q8GT52"
/db_xref="TrEMBL:Q8GT52"
/translation="MSSGAVLVAIAASIGNLGGDNATIAAGVLYIKKEPSLETOPL
TEGLIVAMSLIGATVITTFSGAVADARRPLLIASSVLVFSGLVLMWAPNVVLL
ARLDIGFGLAVTLVLYISETAPTIRGLLNTLPQPSGGGMLSCYMYTWSLMP
QPDRLMGLVSLIPSLMIFALTYPSPRLVSKGRMAEKVLRQLRGREDVSGE
NALIVLEGVKGDTTFEYIIGDPDLADGLADPQEKLYLGAEGVSWTAPRVXG
QSAALGALGMSRHGSMVSKGLVDPLVTLFGSVHEKMPVNGMRSLTFPNFGM
FSVAGQOAKADWDAESHRDDEVDASDHGADIEDNLNPLISQATSVCKEKAAPH
GSRMGVSSSMQGDVAVSMIGGWLAKWTEREGADGQKEGGFQRIYLHEGVGS
GDRGSLIUSFGGDIIPGGEYIQAAALVSQFALYSKDLIEQLAGPAMVHSEAVANG

TKWAELEFPGVKHALFVIGLQILOQFAGINGVLYYTPQILEQAGVGILLNIGLSS
 SASILSALTLLMLESIGIAMLMDSGRRLLSSTIEVLIVAlAILVNVLDVGT
 MHAALSTTSVIVYFFVWGFGPIFNILCABIFPISVRGIGIATCAITFMVLDIIVT
 YTLUPVMLNAILAGLAVGFIYAVVCMIAFVFMVKPBTGKMPLEVITERPVSUAKQKKE
 ATD"

ORIGIN

Alignment Scores:

Pred. No.: 7, 2e-178 Length: 2614
 Score: 2808.00 Matches: 543
 Percent Similarity: 84.39% Conservatives: 84
 Best Local Similarity: 73.08% Mismatches: 104
 Query Match: 68.91% Indels: 12
 DB: 8 Gaps: 6

US-10-051-909-32 (1-800) x HVU534445 (1-2614)

QY 57 SerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAsp 76
 DB 89 TCGGGCGCGTGTGTCGCCATAGCGGGTCCATCGGGAACCTGTCGAGGGTGGAT 148
 QY 77 AsnAlaThrIleAlaAlaValLeuTyrlleLysGluPheGlnLeuGlnAsnGlu 96
 DB 149 AATCGACCATTCAGGTGCGTCTGTACATAAAGAGGAGTTCAGCCCTGGAGACCCAG 208
 QY 97 ProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrThr 116
 DB 209 CCCTGATCAGCGGCTCATCGGCGCATGCTCATCGAGCGACGGTATACAGAGC 268
 QY 117 PheSerGlyProLeuSerAspSerIleGlyArgTrpMetLeuIleLeuSerSerIle 136
 DB 269 TTCCTGGCGCGGTGTCAGCGTGTGTAGCGGCCCTGTCGATCGCTGCTGTC 328
 QY 137 LeuTyPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyPheValLeuLeu 156
 DB 329 CTCCTGTTGTCAGTGGCTGGTGTGTCGTCGCGGCCCAACGTCATGTTGTGCTGTG 388
 QY 157 AlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrlle 176
 DB 389 GCCAGGCTCATCGCGGTTCGGTATCGGTTGGCAGTCACCCCTGTCCTCTGTACATT 448
 QY 177 SerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGly 196
 DB 449 TCGGAGACCGCCCGACTGACATAGAGGCTGCTGAACACACGCTGCGCAGTTCAGTGG 508
 QY 197 SerGlyGlyMetPheLeuSerTyPheCysMetValPheGlyMetSerLeuSerProSer 216
 DB 509 TCAGGAGGAGTGTCTCTTCTTACTGTCATGGTGTTCACCATGTCCTCATGCGCGAGCT 568
 QY 217 AspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhePheGlyLeuThr 236
 DB 569 GACTGGAGAATCATGCTGGGGTTTGTGATCCCGTCCGCTCATGTTTTCATTGACT 628
 QY 237 IlePheTyrlleProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAla 256
 DB 629 GTCCTTTATTTGCCAGTCGCCAGATGCTGTGAGCAAGGAGAAATGCTGAGGCC 688
 QY 257 LysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeuLeu 276
 DB 689 AAGCGAGTATTGCAAGACTCGGGGAGGGAAGATGCTCAGGAGAGATGGCCCTCTT 748
 QY 277 LeuGluGlyLeuGluValGlyAspThrSerIleGluGlyTyrlleIleGlyProAla 296
 DB 749 GTTGAAGGATGGGTGTTGGAAAAGACACATTTTGAGGAATAATATATGACCTGAT 808
 QY 297 ThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrlle 316
 DB 809 GATGACCTGGCTGATGACGGCTGGCTCCA---GATCAAGAGAGTGTAAACTCTATGGA 865
 QY 317 ProGluGluGlnSerTrpIleAlaArgProSerLys-----GlyProIleMetLeu 334
 DB 866 GCTGAAGAGAGGGGTATCTTGGATCGCCCTGCTGTAGGGTGGCGGCCAAAGTGCATT 925

QY 335 GlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeu 354
 DB 926 GGAAGCGCTTTGGGCCTCATGCTTCGTGATGGGAGTATGTTAGTCAGGGTAAATCTCTT 985
 QY 355 MetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGly 374
 DB 986 GTTACCCCTTGTCACTCTCTTCGGAAGTGTCCATGAGAAGATGCTGAGTAAATGGGG 1045
 QY 375 SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHis 394
 DB 1046 AGCATGGGAGCACCTGTTTCTTAACCTTTGGCAGCATGTTAGTGTGGTGAACAGCAG 1105
 QY 395 AlalysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrlleAlaSer 414
 DB 1106 CAGCTAAAGCTGACTGGGATGCTGAGAGT---CATAGGGATGATGAAGATTAATGATCG 1162
 QY 415 AspGlyAlaGlyAspTyrlleGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla 434
 DB 1163 GATCATGCTGATGACATTTGAGGATAACCTCAATAGCCCGCTTATTTCTGTCGAAGCG 1222
 QY 435 ThrGlyAlaGluGlyLysAspIleVal-----HisHisGlyHisArgGlySerAlaLeu 452
 DB 1223 ACAAGCCTGGAGGTAAGGAGATTCTGTCACCTCATGGAAGCATTAATGGTGTGTGGAA 1282
 QY 453 SerMetArgArgGlnThrLeuLeuGlyGluGlyAspGlyValSerSerThrAspIle 472
 DB 1283 AGTAGTAGCATGCAG-----GGAGGGGACCCAGTAGCAGCAGTGGGCATT 1327
 QY 473 GlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLys 492
 DB 1328 GGTGGGGGTGGCAGTTAGCTTGGAACTGACACTGAGAGAGAGTGCAGATGGGCAAAA 1387
 QY 493 GluGlyGlyPheLysArgValTyrlleHisGlnGluGlyValProGlySerArgArgGly 512
 DB 1388 GAAGTGGCTTCAGCGTATTTACTTGCATGAGAGGGTGTGTCAGTGTATCGAGAGGC 1447
 QY 513 SerIleValSerLeuProGlyGlyAspValPheGluGlySerGluPheValHisAla 532
 DB 1448 TCTATATTGCTATGCCA---GGAGGTGATATTCTCTGCTGCTGATATATCCAGGCA 1504
 QY 533 AlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSer 552
 DB 1505 CCGCTCTAGTGACCAACCTGCTCTTACTCGAAGACCTGATAGCAGCAGCTGTCT 1564
 QY 553 AspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeu 572
 DB 1565 GGTCCAGCCATGTACATCCATCCGAGGAGTTCGCCAGGGTACAAAGTGGGCAACTA 1624
 QY 573 PheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnLeuGlnGln 592
 DB 1625 TTTGAACCTGGAGTGAAGCATGCATGTTTGTGGCATTTGATACAGATCTCCGAGCAG 1684
 QY 593 PheAlaGlyIleAsnGlyValLeuTyrlleThrProGlnIleLeuGluAlaGlyVal 612
 DB 1685 TTTGGCGGATCAACGGAGTCTGTATTACACCCCTCAGATCTTTCAGCAGCAGTGT 1744
 QY 613 AlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuSerSer 632
 DB 1745 GGGATTCTTATCAACATTTGACTTAAGCTCTCTCTGCTCTCTTATTATTAGTGC 1804
 QY 633 LeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSer 652
 DB 1805 TTGACAACTTGTGATGCTTCCACATTTGGCATGGCATGCCATGAGACTCATGATATGCA 1864
 QY 653 GlyArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeu 672
 DB 1865 GGAAGAAGGTTCTCTCTCTCTCAACATCCCTGTCTGATAGTGGCACTAGTATCTTG 1924
 QY 673 ValValSerAsnLeuLeuAspLeuGlyThrLeuAlaHisAlaLeuLeuSerValSer 692
 DB 1925 GTTTTGTGAATGTTCTGGATGGGAACCATGGTGCACGGCGCTCTCTCAACAGTCACT 1984
 QY 693 ValIleValTyrllePheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys 712

1985 GTCATAGTCTACTCTGCTTCTTCGTCATGGGTTTGGACCTATCCGGAATATTCTCTGC 2044
713 AlagluilePheProThrArgValArgGlyLeuCysileAlaileCysAlaPheThrPhe 732
2045 GCGGAGATTTCCTCCACCTCCGTCCTGGCATCTGATTGCCATCTGCGCACTAACCTTC 2104
733 TrpIleGlyAspIleValThrThrSerLeuProValMetLeuAsnAlaileGlyLeu 752
2105 TGGATTGGTGACATCATCTGATGACATACACTCTTCGCGTATGCTCAACGCCATCGGACTC 2164
753 AlaglyValPheSerileTyAlaValValCysLeuileSerPheValPheValPheLeu 772
2165 GCTGGAGTTTTCGGAATATATGCTGCTGCTTTCGATGATGACCTTCGATTCTCTACATG 2224
773 LysValProGluThrLysGlyMetProLeuGluValileThrGluPhePheAlaValGly 792
2225 AAGGTCCTGAGACAAAGGCGATGCCCTGGAGTTCATCAGGAGTTCTTCGTCGCGG 2284
793 AlaLysGln 795
2285 GCGAAGCAG 2293

RESULT 8
LOCUS AR208568
DEFINITION Sequence 7 from patent US 6383776.
ACCESSION AR208568
VERSION AR208568.1 GI:21509755
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2601)
AUTHORS Allen, S.M., Hitz, W.D., Kinney, A.J. and Tingey, S.V.
TITLE Plant sugar transport proteins
JOURNAL Patent: US 6383776-A 7 07-MAY-2002;
FEATURES
source 1. 2601
/organism="unknown"
/mol_type="unassigned DNA"

Alignment Scores:
Pred. No.: 6,32e-169 Length: 2601
Score: 2674.00 Matches: 522
Percent Similarity: 82.20% Conservative: 92
Best Local Similarity: 69.88% Mismatches: 115
Query Match: 65.62% Indels: 18
DB: 6 Gaps: 8

US-10-051-909-32 (1-800) x AR208568 (1-2601)

56 MetSerGlyAlaValLeuValAlaileValAlaSerileGlyAsnLeuGlnGlyTrp 75
175 ATGAAGGTCGCTCTGTTGCTATTCGCGCTTCATGTTTCCATGTTGTAATTCCTCCAGGATGG 234
76 AspAsnAlaThrileAlaAlaValLeuTyrlleLysLysGluPheGlnLeuGlnAsn 95
235 GATAATGCTACCATCGCGGGGCTAATGTTTACATTAAGAAAGACCTTGCTTTGGGAACA 294
96 GluProThrValGluGlyLeuileValSerMetSerLeuileGlyAlaThrileValThr 115
295 -----ACTATGAAAGGCTTGTTGGTGGCATCTCCCTGATTGGCAACGCGTAATCACC 348
116 ThrPheSerGlyProLeuSerAspSerileGlyArgProMetLeuileLeuSerSer 135
349 ACTGCTCTGCTCTATAGCGGATTCGCTCGCTCGGCGACCCATGATGATATCTCATCT 408
136 IleLeuTyrrPhePheSerGlyLeuileMetLeuTrpSerProAsnValTyrrValLeuLeu 155
409 GTGCTCTATTCTTGGGTGGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 468

156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrr 175
469 TTGCGGAGGCTACTGTGATGATTTGGGATTTGGGCTTCTGCTGACTCTTGTCCCGGCTCTAT 528
176 IleSerGluileAlaProSerGluileAargGlyLeuLeuAsnThrLeuProGlnPheSer 195
529 ATATCTGAACGCGCGCTCTGAATAAAGGGGCTGTTGAATACGCTTCCTCAGTTCACT 588
196 GlySerGlyMetPheLeuSerTyrrCysMetValPheGlyMetSerLeuSerProSer 215
589 GGCTCTGAGGAATGTTTTCGCTACTGATGTTTTCGCTATGTTTTCGCTATGTTTTCGCTATG 648
216 ProAspTrpArgIleMetLeuGlyValLeuAlaileProSerLeuPhePheGlyLeu 235
649 CTAAGTGGAGGCTCATGCTTGGGCTTCTGCTATTCCTCTCTCTGTTATTTGCAATG 708
236 ThrIlePheTyrrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255
709 ACCATTTTCTTCCCGGAGTCTCTCGGTGGTGGTTCAGCAAGGAAGGATGCTCGAG 768
256 AlaLysLysValLeuGlnLysLeuArgLysAspValSerGlyGluLeuSerLeu 275
769 GCTAAGAGGCTCTCCAAAGATTCCGCGGAAGGAGGATGTTCTCAGCGGAGATGGCATG 828
276 LeuLeuGluGlyLeuGluValGlyAspThrSerileGluGluTyrrileIleGlyPro 295
829 CTGGTTGAAGGCTCTCGGATTTGGGCTGATACATCTATCGAAGAGTACATAAATGGGCT 888
296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrr 315
889 GCTGACGATGTGGCTGATGTCATGAACATGCAACAGAGAAAGATAAAATTCGATTATAT 948
316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335
949 GGATCCCAAGCAGCGCTTCTTGGTTATCAAAACCTGTCACTGCAGACAGATTTCTATTGGC 1008
336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355
1009 -----CTTGGCTCACCATGGAAGAGCATCAACCAAGAGCATGCCCTCATG 1056
356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly----- 373
1057 GATCTCTGGTGACACTGTTTGGTAGCATTCATGAGAGCTCCCGGAGACAGAGCAAGA 1116
374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
1117 GGAAGCATGCAAGCAGCTCTGTTCCAAATTTTGAAGCATGTTTCAAGCATCTGCTGAGCG 1176
394 HisAlaLysAsnGluGlnTrpAspGluAsnLeuHisArgAspAspGluGluTyrrAla 413
1177 CATGCTAAATTTGAACAATGGGATGAAGAAAGCTTACAAGAGGACGCTGAGGACTACATG 1236
414 SerAspGlyAlaGlyGlyAspTyrrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
1237 TCAGATGCAACCCCTGGGAGCTCCGATGATAATTTGCACAGCTCTTAAATCTCAGCCAA 1296
434 AlaThrGlyAlaGluGlyAspIleValHisGlyHisArgGlySerAlaLeu---- 452
1297 ACAACAGCCTTGA-----AAAGACTTACCTCTCTCTCTCCCATGGCAGATCTCTCGC 1353
453 SerMetArgArgGlnThrLeuLeuGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 471
1354 AGCATGAGCGTCAAGTAGTCTCATGCAAGGCTCAGTGAGCAAGGTTGATGATACAGT 1413
472 IleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArg 491
1414 ATTGGTGGTGGCTGCAACTGGCAATGGAAATGGACTGATAAA-----GGTGGAGTGGAAA 1470
492 LysGluGlyGlyPheLysArgValTyrrLeuHisGlnGluGlyValProGlySerArgArg 511
1471 CAACAAGGAGGCTTTAAAGGATTTATTTACATGAGGAGGAGTTTCTGCACTCTCGCT 1530
512 GlySerIleValSerLeuProGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 531

```

Db      1531  GGAATCAATGATCGATCCCGGTGAGGC-----GAAATGTCGAC 1572
Qy      532  AlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMet 551
Db      1573  GCTGCTGCTTGTGTAAGCAACCGCTCTTTACTCCAGGAGCTTATTGATGACACCCA 1632
Qy      552  SerAspAlaValSerValHisProSerGluValAlaAlaValSerLysSerArgTrpLysAsp 571
Db      1633  GTTGGGCGCTGCAATGGTTTACCCATCTGAGACAGCTTCAAGGGCCCAAGTTGGAAGCT 1692
Qy      572  LeuPheGluProGlyValArgAlaLeuLeuValGlyValGlyLeuGlnLeuGln 591
Db      1693  CTTCTTGAACCGGGTTAAGCATGCTGTTGGTTGGAGTTGGAATACAAATCTTCAG 1752
Qy      592  GlnPheAlaGlyLeuGlnGlyValLeuTyrThrProGlnLeuGluGlnAlaGly 611
Db      1753  CAGTTTTCAGGGGATAATGGGGTTCTATATTACACACCTCAATCTTGAAGAGGCCGGT 1812
Qy      612  ValAlaValLeuSerLysPheGlyLeuSerSerAlaSerAlaSerLeuLeuSer 631
Db      1813  GTTGAAGTCTCTTTTCAGATATAGCATTTGGCTCAGATCGGCATCATCTTATCAGT 1872
Qy      632  SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu 651
Db      1873  GCTTTTCACACCTCTTGTGATGCTTCCCTGTATAGGCGTAGCCATGAAGCTCATGATGTT 1932
Qy      652  SerGlyValArgPheLeuLeuLeuGlyThrLeuProIleLeuLeuAlaSerLeuValLeu 671
Db      1933  TCAGGAGAGAGGAGCTGTTACTTACTACAAATCCCGTCTGATTTGTCTCATCTATT 1992
Qy      672  LeuValValSerAsnLeuLeuAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691
Db      1993  TTGGTCAATTCGAAAGCCCTGTAATTTTGGCAATGTCCGCCATCGCAGCAATCTCAACAGTA 2052
Qy      692  SerValLeuValThrPheCysCysPheValMetGlyPheGlyProIleProAsnLeuLeu 711
Db      2053  TGGCTTTGTGTTTATTCTGCTGCTTTGTGATGGTTATGGACCAATTCCAACATCCTT 2112
Qy      712  CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731
Db      2113  TGTCTAGAGATTTTCCCACTAGGCGTGGCTCTGCAATCTGATCTGTCATTAGTG 2172
Qy      732  PheTrpIleGlyAspIleLeuValThrTyrSerLeuProValMetLeuAsnAlaIleGly 751
Db      2173  TTTCTGATTTGGAGACATCATCATCATCTGCTGCTGCTGCTGATCTGCGCTCTTTAGGA 2232
Qy      752  LeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPhe 771
Db      2233  CTTGTTGGTTGATTTCCCATTTACGAGTTGTTTGTTCATCTCGTGATATTTCTGTTT 2292
Qy      772  LeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaVal 791
Db      2293  TTGAAGGTTTCCAGAAACAAAGGCGCATGCCCTTGAAGTCATCTCTGAATCTTTCTGTT 2352
Qy      792  GlyValLysGlnAlaAlaAla 798
Db      2353  GGAGCAAGAGCGCTGCTTCT 2373

RESULT 9
AY094465
LOCUS      2570 bp mRNA linear PLN 05-MAY-2002
DEFINITION Arabidopsis thaliana AT4g35300/F23E12_140 mRNA, complete cds.
ACCESSION AY094465
VERSION AY094465.1 GI:20453188
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
REFERENCE 1 (bases 1 to 2570)
AUTHORS Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J.,

```

Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

Arabidopsis cDNA clones
Unpublished
2 (bases 1 to 2570)

Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

Direct Submission
Submitted (08-APR-2002) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10310 N. Torrey Pines Road, La Jolla, CA 92037, USA

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL CDNA : "RIKEN Arabidopsis Full-length cDNA") : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PCEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L., Chang, E., Dale, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Shinn, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES
source

1..2570
Location/Qualifiers

/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="4"
/clone="RAFL09-48-J19 (R21968)"
/ecotype="Columbia"
209..2428
/note="putative sugar transporter protein"
/codon_start=1
/product="AT4g35300/F23E12_140"
/protein_id="AA019835.1"
/db_xref="GI:20453189"

CDS

ORIGIN

Alignment Scores: 8,48e-168 Length: 2570
Pred. No.: 2657.00 Matches: 530

Db	1297	-----GAAGGTTTCCCGAGATCTCGAGCTGGCTCAATT	1329
Qy	515	ValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla	534
Db	1330	GTTTCAATGCTGGTGTGATGACACCGGTGAG---GCAGATTGTGTACACGCTGCT	1386
Qy	535	LeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAla	554
Db	1387	TTGGTTAGCAACACAGCTCTTTATTCACAAAGACCTTCTCAAAGACATACAAATTGGTCT	1446
Qy	555	AlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGlu	574
Db	1447	GCTATGGTACATCCATCCGAA---ACAACTAAGGTCATTTGGCATGATCTTCATGAT	1503
Qy	575	ProGlyValArgAlaLeuLeuValGlyValGlyLeuGlnLeuGlnPheAla	594
Db	1504	CCTGGAGTCAAGCGGTGATAGTCTAGGAGTTGGACTTCAATAATCTTCAGCAGTTCTCA	1563
Qy	595	GlyLeuGlnValLeuValThrProGlnLeuLeuGlnAlaGlyValAlaVal	614
Db	1564	GGCATCAACGAGTCTTTTACTACACACGCAATCTTGACAGCGGGTCTCGGATC	1623
Qy	615	IleLeuSerLysPheGlyLeuSerSerAlaSerIleLeuLeuSerSerLeuThr	634
Db	1624	CTACTATCGAATCGGATAGTCTTCTCAGCATCTCTTACTATAAGTGCATTGACA	1683
Qy	635	ThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg	654
Db	1684	ACCTTTGTGATGTACTCGCAATAGCTGTGGCAATGAGGCTCATGATCTTCTGTGCGA	1743
Qy	655	ArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValLeuVal	674
Db	1744	AGACCTTGCTTCTCACCAGATACCAATCTCTAGTATCTCTATTGGTTTAGTAATC	1803
Qy	675	SerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIle	694
Db	1804	TCAAATCTTGTTCATGACACACATCTGTCACGCGGTCTTATCAACCGTAAGCGTTGTG	1863
Qy	695	ValThrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlu	714
Db	1864	CTCTACTTCTGCTTCTGATGAGGTTTGGTCTGCTCTCAACATCTCTCTGTTCCAGAG	1923
Qy	715	IlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIle	734
Db	1924	ATTTTTCCAACTCGAGTCCGCGGAATCTGCATCGCCATCTGCGCATCTACCTTCTGGATC	1983
Qy	735	GlyAspIleIleValThrValSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGly	754
Db	1984	TGTGACATAATCGTCATTCAGTCTCCCGTGTCTCAATCCATTGACATGCTGCT	2043
Qy	755	ValPheSerIleValAlaValValCysLeuIleSerPheValPheValPheLeuLysVal	774
Db	2044	GTGTTTGAATGTACGCAATCGATGTGCTATTTTCATGGTCTTTGTGTTCAATAAGTC	2103
Qy	775	ProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValAlaLys	794
Db	2104	CCGAAATCAAGGCATGCCATTTGAATGATCATCAGATCTCTTTCTGTGGAGCTAGA	2163
Qy	795	GlnAlaAlaAla 798	
Db	2164	CAAGCTGAGCT 2175	
RESULT 11			
LOCUS	ATH532570	2190 bp	linear
DEFINITION	Arabidopsis thaliana mRNA for monosaccharide sensing protein 2		
ACCESSION	AJ532570		
VERSION	AJ532570.1	GI:26800694	
KEYWORDS	monosaccharide sensing protein 2; mssp2 gene.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE			
AUTHORS	Stamme, C., Tjaden, J., Trentmann, O., Emmerlich, V. and Neuhaus E.		
TITLE	A novel family of plant monosaccharide transporters is involved in a new type of eukaryotic sugar sensing		
JOURNAL	Unpublished		
AUTHORS	Neuhaus, E.		
TITLE	Direct Submission		
JOURNAL	Submitted (04-DEC-2002) Neuhaus E., Plant Physiology, University of Kaiserslautern, Erwin-Schroedinger-Str. 22, 67663 Kaiserslautern, GERMANY		
FEATURES			
source	1. .2190		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/db_xref="taxon:3702"		
gene	1. .2190		
	/gene="mssp2"		
CDS	1. .2190		
	/gene="mssp2"		
	/codon_start=1		
	/product="monosaccharide sensing protein 2"		
	/protein_id="CAD58692.1"		
	/db_xref="GI:26800695"		
	/db_xref="GOA:O65497"		
	/db_xref="TrEMBL:O65497"		
	/translation="MSGAVLVAAIAAVNLIQGMNATAGAVLYIKKEFNLESPSV EGLIVMSLIGATLITTCGSGVADPGRPMILSSILYFVGLWMLSNPNVYVLLLG RLLDGFVGLVTLVPIVYISETAPPEIRGLLNTLPQFTGGGMFLSYCMVGMILMP SPSRMILGVLFIPLVFLFVFLPSPRLVSKGRMLEAKRVQLRGREDVSGEM ALLVEGLIGGETTIEYIIIGPADEVDDDDHDAVDKQIKLYGAEGLSWVAPVKGG STMSVLSRHGSTMRSRQSLIDPLVTLFSGVHEKMPDTGSMRSALFPFGSNFVGGN QPRHEDENLVGEGEDYPSDHGDDSDDDLHSLPLIQRQTTSMKMDPHHTAHTLSTF RHGSQVQAGGAGAGMGIGGQVQVAVKWTREDESGKRGEPGSRGRTSVLPFGD TGGADVQASALYSQALYSKLLKEHTIGPAMVHPSETTKGSIWHDLPDQVKRAL WVGVLQIQPFGSINGVLYYTPQILFQAGVGLLSNMGISSSASLILISALTTFVML PAIAVKMLDLSGRRLTLTLFILASLLVLSNLVHNSIVHVALSVSVLYLF CFVWFGPAPNLCSEIFFRVRGICAIICALTFWICDIIITVTSPLVSLKSLAGLV FGMAIVCCISWVFFIKVPETKGMPLVITEFFSVGARGAARNA"		
ORIGIN			
Alignment Scores:			
Pred. No.:	5,04e-163	Length:	2190
Score:	2584.00	Matches:	512
Percent Similarity:	80.78%	Conservative:	59
Best Local Similarity:	68.82%	Mismatches:	123
Query Match:	63.41%	Indels:	20
DB:	8	Gaps:	8
US-10-051-909-32 (1-800) x ATH532570 (1-2190)			
Qy	55	MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp	75
Db	1	ATGAGTGGAGCTGTCTTGTCTATTGCTGCTGCTGTGCAACTTGTACAGGATGG	60
Qy	76	AspAsnAlaThrIleAlaAlaValLeuValIleLysGluPheGlnLeuGlnAsn	95
Db	61	GATAACGCACTATTTCAGGAGCTGTGTGTATACATAAAGAGATTTAATTGGAGAT	120
Qy	96	GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr	115
Db	121	AATCCATCAGTGAAGGCTAATTGCGGATGTCACTTATTTGGTCTACTGTATTACA	180
Qy	116	ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSer	135
Db	181	ACATGCTCTGGAGGGTAGCTGATTGGTGTGTCGCCGCTCCCATGCTAATATGTCCTA	240
Qy	136	IleLeuThrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTrpValLeuLeu	155
Db	241	ATTCTCTACTTGTGCTTCTCTAGTAAGCTATGGTCTCCGAAAGTTTATGTTGCTC	300

Qy 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
 Db 301 TTAGGAAGGTGTAGATGGATTGGGGTGGCTTTGGTGCACACTGTGTCCTATTAT 360
 Qy 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
 Db 361 ATATCTGAGACTGCACCACTGAGATTAGGGACTGTTGAATACGCTACCGCAGTTCAC 420
 Qy 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
 Db 421 GGCTCTGGAGGATGTCTTATCTTACTGATGATGGTTTCGGAATGTCGTTGATGCCATCA 480
 Qy 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
 Db 481 CCTAGCTGGAGATTGATGCTGTGGTCCCTTTTCATCCCTTCCTTCTTCTTCTTCCTC 540
 Qy 236 ThrIlePheTyrLeuProGlnSerProArgTrpLeuValSerIleGlyArgMetAlaGlu 255
 Db 541 ACAGCTCTCTTCTTGGCCGAGTCCCAAGAGTGCCTGTGAGCAAGGTGCAATGCTTGA 600
 Qy 256 AlaIleGlyValLeuGlnIleGlyLeuArgGlyLeuAspValSerGlyLeuSerLeu 275
 Db 601 GCAAGCGGGTCTTCTCAGAGCTGCGTGGTCGCAAGATGTGTCTGGTGGAGTGGCTTG 660
 Qy 276 LeuLeuGlyLeuGluValGlyGlyAspThrSerIleGluIleGlyPro 295
 Db 661 TTGGTTGAGGCTCTTGGAAATGGAGTGAACCAACCATAGAGGAATATATATTTGGTCCC 720
 Qy 296 AlaThrGluAlaAlaAspLeuValThrAspGlyAspGlyGluGlnIleThrLeuTyr 315
 Db 721 CGCGATGAAGTTACTGATCATGATATAGTGTGCGATAAGGATCAAAATTAAGTTATAT 780
 Qy 316 GlyProGluGlyGlnSerTrpIleAlaArgProSerLeuGlyProIleMetLeuGly 335
 Db 781 GTGCGAGAAGAGGCTGAGTTGGTGGTGTAGCCAGTCAAGGA-----GGA 828
 Qy 336 SerValLeuSerLeuAlaSerArgHisGlySer---MetValAsnGlnSerValProLeu 354
 Db 829 AGCACTATGATGTTTGTCTGCCATGGAAGTACAATGAGCAGGAGGCAAGGCTCATTTG 888
 Qy 355 MetAspProIleValThrPheGlySerValHisGluAsnMetProGlnIleGlyGly 374
 Db 889 ATTGATCCTCTTGTACACTGTTTGGAGCGTTCAGAGAAATGCCGGACACT---GGA 945
 Qy 375 SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHis 394
 Db 946 AGCATGAGGAGTCCCTTGTCCACATTTTGGAGTATGTTCCAGTTTGGAGGAATCAA 1005
 Qy 395 AlaIleAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSer 414
 Db 1006 CCAAGCATCAAGATTGGGATGAGAGAACTTGTGGAGAAGGTGAGGATATATCCATCC 1065
 Qy 415 AspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla 434
 Db 1066 GAC---CATGGAGATGATTTCTGAGATGATCTTCACTTCCGTTGATCTCAGTCAAACG 1122
 Qy 435 ThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMet 454
 Db 1123 ACAAGCATGGAG---AAAGACATGCTCAGCTGCTCATGTAACCTTCTTCTACCTTCAGA 1179
 Qy 455 ArgArgGlnThrLeuLeuGlyGluGlyAspGlyValSerSerThrAspIleGlyGly 474
 Db 1180 CATGGAGTCAAGTCAGGAGCTCAAGGGAGAGGAGCGGTAGTATCGGGATTGGAGGT 1239
 Qy 475 GlyTrpGlnLeuAlaTrpIleTrpSerGluIleGlyGluAsnGlyArgGlyGlyGly 494
 Db 1240 GGATGGCAAGTGGCATGGAATGGAGGAGGAGAGAGATGAATCGGACAGAGAA--- 1296
 Qy 495 GlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgGlySerIle 514
 Db 1297 -----GAAGGTTTCCAGGATCTCGACGTTGGCTCAATT 1329

Qy 515 ValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaAala 534
 Db 1330 GTTTTCATTGCTGGTGGTGTGATGGAAACCGGTGAG---GCAGATTTTGTACAGCGCTGCT 1386
 Qy 535 LeuValSerGlnSerAlaLeuPheSerIleGlyLeuAlaGluProArgMetSerAspAla 554
 Db 1387 TTGGTTAGCCCAACAGACTCTTTATTCCAAGACCTTCTCAAGACATACATAATTTGTCCT 1446
 Qy 555 AlaMetValHisProSerGluValAlaAlaIleGlySerArgTrpIleAspLeuPheGlu 574
 Db 1447 GCTATGGTACATCCATCCGAA---ACAACCTAAGGGTCAATTTGGCATGATCTTCATGAT 1503
 Qy 575 ProGlyValArgAlaLeuValGlyValGlyIleGlnIleLeuGlnIleGlnPheAla 594
 Db 1504 CCTGGAGTCAAGCGTGCATTTAGTCTAGGAGTTGGACTTCAATATCTTCAGCAGTTCTCA 1563
 Qy 595 GlyIleAsnGlyValLeuTyrThrProGlnIleLeuGluGlnAlaGlyValAlaVal 614
 Db 1564 GGCATCAACGAGTCTTTACTACACACCGCAAACTCTTCAGCAGCGGGTGTGGGATC 1623
 Qy 615 IleLeuSerIlePheGlyLeuSerSerAlaSerAlaSerIleLeuLeuSerSerLeuThr 634
 Db 1624 CTACTATCGAACATGGGGATTTCTTCTCAGCATCTTACTTATTAAGTGCATTGACA 1683
 Qy 635 ThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg 654
 Db 1684 ACCTTTGTGATGTACTCGCAATAGCTGTTGCAATGAGGCTCATGGATCTTTCTGGTGA 1743
 Qy 655 ArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValVal 674
 Db 1744 AGGACCTTGTCTTCTCACACGATACCAATCTCTAGCATCTCTATTGTTTGAATAATC 1803
 Qy 675 SerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIle 694
 Db 1804 TCAATCTTGTTCATGAACAGCATTTGTCAGCGGTCTTATCAACCGTAAAGCTTTG 1863
 Qy 695 ValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlu 714
 Db 1864 CTCTACTTCTGCTTCTGCTGATGGGTTTCGGTCTGCTCCAAACATCTCTCTGTCAGAG 1923
 Qy 715 IlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheIle 734
 Db 1924 ATTTTCCAACTCGAGTCCGCGAACTGTCATCGCATCTGCGCATCTACCTTCTGATC 1983
 Qy 735 GlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGly 754
 Db 1984 TGTGACATAATCTGCTACTTACAGTCTCCCGCTGCTGCTCAAAATCCATTGGAGTGTGT 2043
 Qy 755 ValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysVal 774
 Db 2044 GTGTTTGAATGATGACCAATCGTATGTCATTTTCATGGTCTTGTGTTTCAATAAGTC 2103
 Qy 775 ProGluThrIleGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAlaLys 794
 Db 2104 CCGGAACCTAAAGCATGCCACTTGAATCATCACAGAGTCTCTTCTGTTGGAGCTAGA 2163
 Qy 795 GlnAlaAlaAla 798
 Db 2164 CAAGCTGAAGCT 2175
 RESULT 12
 AK065191
 LOCUS
 DEFINITION
 Oriza sativa (japonica cultivar-group) cDNA clone:J013002E10, full insert sequence.
 AK065191
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Oriza sativa (japonica cultivar-group)
 Oriza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
AUTHORS

1 The Rice Full-length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-length cDNA Project Team,
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
Onoe, K., Shishiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group: Otono, Y., Murakami, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oeato, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M., and Hayashizaki, Y.

TITLE
japonica rice
Collection, mapping, and annotation of over 28,000 cDNA clones from
Science 301 (5631), 376-379 (2003)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

2 (bases 1 to 2368)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,
Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
Ootani, N., Ota, Y., Otono, Y., Satoh, H., Sakai, C., Sakai, K.,
Sakazume, N., Sato, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
Yanada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
Yoshimura, A.

Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,
Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica
rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and
Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otono, Y., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Imotani, K.,
Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,

Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Oeato, N.,
Ota, Y., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A. and Hayashizaki, Y.

FEATURES
source

Location/Qualifiers
1..2368
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J013002E10"

ORIGIN

Alignment Scores:

Pred. No.: 6,58e-163 Length: 2368
Score: 2583.00 Matches: 508
Percent Similarity: 83.74% Conservative: 69
Best Local Similarity: 73.73% Mismatches: 106
Query Match: 63.39% Indels: 7
DB: 8 Gaps: 5

US-10-051-909-32 (1-800) x AK065191 (1-2368)

QY 109 IleglyAlaThrIleValThrPheSerGlyProLeuSerAspSerIleGlyA-gAsg 128
DB 2 CTTGGCGGCGAGCATCATCACGAGCTTCTTGGAGCAGTGGCTGATCTTTTGGTGGCGG 61
QY 129 ProMetLeuLeuLeuSerSerIleLeuTyPhePheSerGlyLeuLeuMetLeuTrpSer 148
DB 62 CCCATGCTGATCGCGTGGCTGCTCTACTTTTGTAGTGGCTAGTGTCTTGGCGG 121
QY 149 ProAsnValTyValLeuLeuLeuAlaA-gPheValAspGlyPheGlyIleGlyLeuAla 168
DB 122 CCAATGTGATGTGTGCTCTTGGCGAGGCTCATTTGACGGGTTCGGGATCGGTTTGGCT 181
QY 169 ValThrLeuValProLeuTyTrIleSerGluIleAlaProSerGluIleArgGlyLeuLeu 188
DB 182 GTCACGGTGTGACCATTTGTATCTCTGTGACATGCGCCGACGACATCAGAGACTGCTA 241
QY 189 AsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeuSerTyTyCysMetValPhe 208
DB 242 AACACGGTGGCGGAGTTCAGTGGGTCTGGAGGAGATGTTCTTTCATCATCTGCATGTTT 301
QY 209 GlyMetSerLeuSerProSerProAspTrpArgIleMetLeuGlyValLeuAlaIlePro 228
DB 302 GGCAATGCTCTATGCCACAGCCAGATTTGAGGATCATGCTTGGCGTTCATCAATACCA 361
QY 229 SerLeuPhePhePheGlyLeuThrIlePheTyLeuProGluSerProArgTrpLeuVal 248
DB 362 TCATTATATACCTTGCATTGACCATCTTTTACTTACCTGATCCGCCGAGGTGGCTCGTG 421
QY 249 SerTyGlyArgMetAlaGluAlaLysValLeuGlnLysLeuArgGlyLysAspAsp 268
DB 422 AGCAAGGAAGAATGGGTGAGGCGCAACGCTGTGTTCAAGGCGCTGCGTGAAGAGAAGAT 481
QY 269 ValSerGlyGluLeuSerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIle 288
DB 482 GTTTCAGAGAAATGGCCCTTCTCGTTGAGGCTCTGGGGTTCGGGAAACACACAAAT 541
QY 289 GluGluTyTrIleLeGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAsp 308
DB 542 GAGGATACATAAATTTGGACCTGATGATGAGCTTGTGCTGATCAAGGCGCTCCA---GAT 598
QY 309 LysGluGlnIleThrLeuTyGlyProGluGluGlyGlnSerTrpIleAlaArgProSer 328
DB 599 CCAGAGAGATCAACCTGATGCTCTCTGAGAGAGGCTTATCGTGGGTTCGCCGTCTGTT 658
QY 329 LysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetVal 348
DB 659 CACGGGCAAGTGCATCTGGAGTGCATAGGTCTCATCTCTCGTTCATCGTGTAGTAGTGC 718

annotated as hypothetical proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are identified by RepeatMasker (Arian Smit, <http://ftp.genome.washington.edu/RX/RepeatMasker.html>).

FEATURES

source
1. .142114
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:39947"
/chromosome="10"
/map="RZ500"
/clone="OSJNB0064P21.1"
complement (3008..3658)
/gene="OSJNB0064P21.16"
/note="predicted by genscan"
complement (<3008..>3658)
/gene="OSJNB0064P21.16"
complement (3008..3658)
/gene="OSJNB0064P21.16"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAG46103.1"
/db_xref="GI:12039315"
/translation="MVAGDSQARLLVIALRLLLDPFAAAAAPELFRHSDYRATVP
ARGISVDFVWAPESNLPLMHEDLFLAPRTDVLVIGSLWMLHVTDAARYGALLA
SVVDAKSLRSLPVPVPHMFWLGLPLLVNMLNTDAKKVHNDTLOAYDLEVRG
LQRDGPPLLLDVGLSRGCGQCTADGWHYDGDVYDAVLHMLNALVIESQRI"
complement (3574..3596)
/repeat_region
/rpt_family="GC_rich"
complement (4269..4297)
/rpt_family="(GGGA)n"
5613..5633
/rpt_family="(GGAA)n"
complement (5634..6005)
/gene="OSJNB0064P21.1"
/note="identical to histone H4 (H4C13) GI:169499 (Zea
mays), (Plant Physiol Biochem 25, 235 247 (1987));EST
D23040 from this gene"
complement (<5634..>6005)
/gene="OSJNB0064P21.1"
complement (5634..5945)
/gene="OSJNB0064P21.1"
/codon_start=1
/product="histone H4"
/protein_id="AAG46106.1"
/db_xref="GI:12039318"
/translation="MSGRGKGKGLGKGAKHRKVLRDNIQITKPAIRRLARRGV
KRISGLIYETRGVLFIENVRDATTTEHARRKVTAMDVVALKEQRTLYRG
G"
complement (7473..7546)
/rpt_family="(CGG)n"
7556..7630
/rpt_family="(GGAGAA)n"
complement (8238..8259)
/rpt_family="AT_rich"
complement (8915..8963)
/rpt_family="AT_rich"
complement (11055..11097)
/rpt_family="AT_rich"
complement (11244..11264)
/rpt_family="AT_rich"
14313..14336
/gene="OSJNB0064P21.2"
/note="identical to calcium dependent protein kinase
GI:435466 (Oryza sativa) (Gene 129 (2), 183 189 (1993));
EST C22538, C22540, AA750452, C99271, C22539 from this
gene"
join(14313..15004,15338..15481,15564..15716,15894..16009,
16446..16810,17112..17333,17903..18336)
/gene="OSJNB0064P21.2"
misc_feature
16446..16810,17112..17333,17903..18336
/note="with dinucleotide (AT) repeat (35 to 42 copies)"
34449..38345
/gene="OSJNB0064P21.3"
/note="similar to sugar transporter GI:1495273
(Arabidopsis thaliana); EST AA231788, D48171, C72446,
/gene="OSJNB0064P21.2"

CDS

join(14323..15004,15338..15481,15564..15716,15894..16009,
16446..16810,17112..17333,17903..18025)
/gene="OSJNB0064P21.2"
/codon_start=1
/product="calcium-dependent protein kinase"
/protein_id="AAG46110.1"
/db_xref="GI:12039322"
/translation="MGNSCQGTGYNNYQNSRNFQDRFASRYVDGNDTECVSGSSR
ASLAKRQGLNLSKPVLYGKTNVRLYTLGRLEGQFGKTVLCTEISTGCOYACK
TILKSNLRQVSDIEDVRREIQIMHLSGQKNIVTIKDYDEQAQVHIVMELCAGGELF
SKTKRGHYSERKAAELIKIIGIETCSHGVMHRLDKPENFLLDADDEFVSVAID
FGLSVFRPQOVREVGVSPYIYTAPEVLEXYRGYEPADINTAGVILYLLGPPFPWAD
TQSGIYBKVDGRIDFKSNRWPRISDAKDLIKMLCPYSERLKAHEVLKHPICDN
GVATRALDPVLPLKQFSAMRLKLSIQIIABRLUSEBEI VGLRNFMKAMDTKNS
VVTFGELKGLKRSYVFKDEINDMEADDTTSTINWEEFI AAASVLSNIEKFKHLM
AAFTYFDKDGSGFITVDKLOKACMERNMEDTLEEMILEVDQNDGQIDYAEFVTMQ
SNNFGLQWTVESLVALREAPQVY"
complement (18675..18732)
/rpt_family="GC_rich"
complement (20605..20634)
/rpt_family="AT_rich"
22393..24891
/gene="OSJNB0064P21.6"
/note="predicted by genemark hmm"
complement (22393..24891)
/gene="OSJNB0064P21.6"
22393..24891
/gene="OSJNB0064P21.6"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAG46113.1"
/db_xref="GI:12039325"
/translation="MAGSGSEAMASSGLDIGELDRDPDAIHRRAASVAIVRSGGDGP
RIVDGRDRRRARSRRLSUSSWRPAGGETSVESRRYGTFAAAAEEVGGKTRAAAIW
EMKPVRLSIRIGRERCGLLSVEVAGVGVPSMDGLRLAVTVKRAETKDGAVOTMPA
TVRGCADEFVETLFVRNIYFAGGATGKPLKLEPRFVSVVPAEAPVGLGAHVD
VSLVDSLOKSSSEGRVRWFDATVILSGKATGGLLKLGLQMDAGLCGLYQAAT
EKVDVDDVSPALARAHNKNSFSVARTSGPKLSASDAAI3PSMEAYKQLIDRLSVDEH
GDVTSLLPRKLADDELSDVGLPEYEVVKGVETKEVYVYHAHRDVLKELDSIAEQ
TEALMTNGKKSPSPKPKYDQKQCLDDEEMTVTFILKLEVDKGRKLQKQPT
PRSEKAAAAPVVDLGLGTAVTQDGGFLVSMFPFDLPLEKGAOPPKLAMQ
VSRPFLPSSMAATFVLOKAAAGGADEVDRKVLGGMDNLTGKTPEQVPEGIA
EAVTGGRTGAGSASRAVRLVKLAAAVSDGSESRVATGIMTAADDPTLEVIAP
SLOKLEMAVDALMIQAEADADAPFEVAPAGDATTVDALVSPDSWSESRSQDGRV
TVAAIOLRDPSSREYAVGAPMVAVOSARLIGAAGNSGGRFKVRSLHVGQVQURCPV
CGGAGGGASWGAERQKLTANQWALAHGPARAAGERARTPTTTTTPSSQARQRPDVW
SLSSRVLAGMMLKTVRNPDVVRVGTAAAGSGGGN"
complement (23766..23788)
/rpt_family="GC_rich"
complement (24616..24639)
/rpt_family="GC_rich"
complement (24693..24714)
/rpt_family="GC_rich"
24853..24903
/rpt_family="(CGG)n"
complement (26149..26194)
/rpt_family="AT_rich"
complement (26861..26905)
/rpt_family="AT_rich"
complement (27502..27537)
/rpt_family="GC_rich"
complement (27628..27663)
/rpt_family="AT_rich"
complement (28100..28132)
/rpt_family="AT_rich"
complement (29892..29929)
/rpt_family="(TA)n"
29983..29928
/note="with dinucleotide (AT) repeat (35 to 42 copies)"
34449..38345
/gene="OSJNB0064P21.3"
/note="similar to sugar transporter GI:1495273
(Arabidopsis thaliana); EST AA231788, D48171, C72446,
/gene="OSJNB0064P21.2"

repeat_region

repeat_region

gene

mRNA

CDS

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

Db	36781	CTAAGGAATTGATGGAGCAACGCCCTTCTGGCCCTCTGCTATGCTGCATCCATCTCAGGCAG	36840
Qy	563	laAlaLysGlySerArgTTPtLysAspLeuPheGluProGlyValArgArgAlaLeuLeuV	583
Db	36841	TTGCTAAAGTCCAAATGGCAGACTTATCGAACCTGGAGTGAACATGCTCTGTGG	36900
Qy	583	alGlyValGlyIleGlnIleLeuGln-	591
Db	36901	TTGGCATAGGATACAAATCTCGAACAGGTAAACCTTTTCTGATTTTATGTCATGTAC	36960
Qy	591	-----	591
Db	36961	GAATGATGTGCTATGGTTGAACCTTAATTTCTAGTAGCTCCAGTATAAATTACTCTGAT	37020
Qy	591	-----	591
Db	37021	TAATGTAACGTGTTAAAAATCAAGACATCTTCATGCAGTCTCTCTAGGTTCTTAGCGAG	37080
Qy	591	-----	591
Db	37081	TCCTCATTAAGACTCTTACTTACCTTACTCGAAAAACATGTTTAGCTGTGTACGTGTTT	37140
Qy	591	-----	591
Db	37141	TCITGCATATATTCATACACAGGACATGAGCTAAACCTGCACAAATAATACACCTTTG	37200
Qy	591	-----	591
Db	37201	ATGATTAGAAATCAAAATGCCATCAATGTTGTTTGTAGTTTACACTAGTAAGTCAACCT	37260
Qy	591	-----	591
Db	37261	TAACACTTCACAGTATTCATTATTAGAAAAACCATGGGACTGGGCTATGTCATGACACC	37320
Qy	591	-----	591
Db	37321	TGATTACGGTATTCATCTGAAGCAGAAAAATGTCATGCATTCATTTTATTTCCTG	37380
Qy	591	-----	591
Db	37381	AATTTATTGCTGAATATGGTCTAGTTTCGGCAGAAATAACTTTCAGAAACTTTTCACAGT	37440
Qy	591	-----	591
Db	37441	GGCATAATCATGATATTCCTTTTACTTCCCTCCACTCTCTATTTCCCTTGTGAT	37500
Qy	591	-----	591
Db	37501	ATTTTATATCAGTCTATTATGTTTACCTCATAGCATACTGAATGTTAATGTGCTAT	37560
Qy	592	-----GlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGln	605
Db	37561	CAACTTTTCCATTATGTCAGTTTGTGGCAITTAATGGAGTTCTGTACTACATCCACAA	37620
Qy	606	IleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSer	625
Db	37621	ATTCCTGAGCAAGCTGGTGTGTGTCTCTTTCGAAACATTTGACTAGTCCCTCATCT	37680
Qy	626	AlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAla	645
Db	37681	GCACTATTCTTATTACGCACTGACAACTTGCTGATCTCTCCAGCATTTGGTATTGCT	37740
Qy	646	MetLeuLeuMetAspLeuSerGlyArgArg-	655
Db	37741	ATGAGGCTCATGGATATGTCGGAAGAAGTTAGACTTCATATTTTACATTTTTTTTGT	37800
Qy	655	-----	655
Db	37801	GATTACATGCATCACTCTTAAGTGAAGATATATGCTAACCATATTATCTGGCTGTCTCT	37860
Qy	656	-----PheLeuLeuGlyThrIleProIleIleAlaSerLeuValIleLeu	672

Db	37861	CATTCAAGTTTCCTTCTCCTTCGCAACATCCCTATCTCTGTATAGTAGACTAGTATCTTG	37919
Qy	673	ValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValser	692
Db	37920	ATTCTGGTCAATATTCTCGATGTGGGACCATGGTTCATGCCTCACTGTCCACAGTCACT	37979
Qy	693	ValIleValTyPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys	712
Db	37980	GTCATACTCTACTTCGTCTTTGTTCATGGGGTTCGGGCCCTATTCCAAACATCTCTGT	38039
Qy	713	AlaGluIlePheProThrArgValargGlyLeuCysIleAlaIleCysAlaPheThrPhe	732
Db	38040	GCAGAGATTTCOCGACCAACCGTGTGGCATCTGCATAGGCATCTGTGCCTAACATTC	38099
Qy	733	TpIleGlyAspIleIleValThrTyrserIeuProValMetLeuAsnAlaIleGlyLeu	752
Db	38100	TGATCGGTGATATCATTTGATACATACACCCCTCCCCTGGATGCTCAACGCCATTGGACTC	38159
Qy	753	AlaGlyValPheSerIleTyraLaValValCysLysLeuIleSerPheValPheValPheLeu	772
Db	38160	GCTGGAGTGTTCGATCTACCGAGTGGTCTGCATACTGGCTTTCCTGTTGCTTCATG	38219
Qy	773	LysValProGluThryrGlyMetProLeuGluHuValIleThrGluPheAlaValGly	792
Db	38220	AAGGTCCGGAGACAAGGCGCATGCTCTTGAAGTCATCACCGAGTTCTTCTCTGTGCGA	38279
Qy	793	AlalyscIlnAla	796
Db	38280	GCAAAGCAGGCC	38291
RESULT 14			
AE017116			
LOCUS	Oryza sativa (japonica cultivar-group)	DNA linear	PLN 06-JUN-2003
DEFINITION	77 of the complete sequence.		chromosome 10, section 70 of
ACCESSION	AE017116		
VERSION	AE017116.1		
KEYWORDS	GI:31433303		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
AUTHORS	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
CONSTRM	Eurhartoideae; Oryzedeae; Oryza.		
TITLE	1. (bases 1 to 300957)		
JOURNAL	The Rice Chromosome 10 Sequencing Consortium		
REFERENCE	In-depth view of structure, activity, and evolution of rice chromosome 10		
AUTHORS	Science 300, 1566-1569 (2003)		
CONSTRM	2 (bases 1 to 300957)		
TITLE	Buel,C.R., Wing,R.A., McCombie,W.R., Messing,J. and Yuan,Q.		
JOURNAL	Submitted (05-MAY-2003) The Institute for Genomic Research, 9712		
AUTHORS	Medical Center Dr, Rockville, MD 20850, USA		
COMMENT	This is the pseudomolecule for rice chromosome 10, which was constructed by resolving discrepancies between overlapping BACs, trimming the overlap regions, and linking the unique sequences to form a contiguous sequence. Genes in individual BAC clone were identified by a combination of several methods: Gene prediction programs, searches of the complete sequence against a peptide database and EST databases. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as unknown proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as hypothetical proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).		
FEATURES	Location/Qualifiers		
SOURCE	1..300957		
	/organism="Oryza sativa (japonica cultivar-group)"		
	/mol_type="genomic DNA"		
	/cultivar="Nipponbare"		

29890..30054,30356..30577,31147..31580)
/locus_tag="OSUNB0064p21.2"
join(27567..28248,28582..28725,28808..28960,29138..29253,
29890..30054,30356..30577,31147..31269)
/locus_tag="OSUNB0064p21.2"
/codon_start=1
/product="calcium-dependent protein kinase"
/protein_id="AAP54840.1"
/db_xref="GI:31433311"
/translation="MGNSSQNGTYGNNYQNSRNFQNDRFASRYVGNEDTDCYSSSR
ASLAGRLGKLNKSPVLGYKTPNVRELTLRGQGFQKTYLCTELSTCCQVACK
TLKSLRVCSDIDVERRIQIMHLSGKRNIVTKDTEQAVHIVMELCAGELF
SKIQKRGHSYERKAELIKIIVGIIETCHSHGVHRDLKPENFLILLADDEFSVRAID

Alignment Scores:

Pred. No.: 2,96e-155 Length: 300957
Score: 2514.00 Matches: 552
Percent Similarity: 53.77% Conservative: 75
Best Local Similarity: 47.34% Mismatches: 109
Query Match: 61.69% Indels: 431
DB: 8 Gaps: 9

US-10-051-909-32 (1-800) x AB017116 (1-300957)

Qy 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuGlnGlyTyr 75
Db 48060 ATGGCGGCGCGTGTGTGCGCATCGCGGCTCCATCGCAACTTGTCTGAGGCTGG 48119
Qy 76 AspAsnAlaThrIleAla----- 81
Db 48120 GATAATGCAACCAATTCGAGTAAGTTCGTGGTTAGAGAAAAACGAATCTTTTTTTTCT 48179
Qy 81 ----- 81
Db 48180 GCCATAGTTTGTGCGCATCTGTATGTGACGCTAGGCGCGTGTGTGTCGCAATTAT 48239
Qy 81 ----- 81
Db 48240 TGTGATACTTTGTGAGATCATGATGATTTGGTGAATCATCTGCAGTAAAAATAGTTT 48299
Qy 81 ----- 81
Db 48300 AGTTGTCTTCTGTTCATGATATCAAAAAAACAACACTGATATTTTTTTCAGTA 48359
Qy 81 ----- 81
Db 48360 GAATTTTGTGATCAATTATGAGTAGGATTTAACTAATGTACGGATCTCATGAATGATCTGT 48419
Qy 81 ----- 81
Db 48420 ATTTATGGAACTTGATATCTGTGCAATCGTAAGTTTCAGAAATTTGATGATACCGATG 48479
Qy 82 -----AlaAlaValle 85
Db 48480 GTTTTGACTTCTGCTGACCTCGGGACTTGTGGGATGATCTTTATCAGGTGCGGTACT 48539
Qy 85 uTyrIleGlyGluPheGlnLeuGlnAsnGluProThrValGluGlyLeuIleValse 105
Db 48540 GTACATCAAGAAGAAATTCACATTGCGAGCGAGCCCTTATCAAGGCCGTGATCGTGGC 48599
Qy 105 rMetSerLeuIleGlyAlaThrIleValThrPheSerGlyProLeuSerAspSeril 125
Db 48600 CATGTGCTCATTTGGGCGGACGATCATCAGCGTCTCTGAGCAGTGGCTGATCTTT 48659
Qy 125 eGlyArgProMetLeuLeuLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMe 145
Db 48660 TGGTAGGCGGCCCATGTGATCGGTGCGGTCTCTCTACTTTGTTAGTGGGTAGTGAT 48719
Qy 145 tLeuTrpSerProAsnValTyrValLeuLeuAlaArgPheValAspGlyPheGlyI 165
Db 48720 GCTTTGGCGGCAATGTGATGTTGCTCTTGGCGAGGCTCATTCACCGGTTGGGAT 48779
Qy 165 eGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaProSerGluIleAr 185

Db 48780 CGGTTTGGCTGTCAAGCTGTGTACCATCTCTGAGACTGCCCGACGACATCAG 48839
Qy 185 gGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMetPheLeuSerTyrCy 205
Db 48840 AGGACTGTAAACACGCTGCCGCGAGTTCAGTGGTCTCTGGAGGGATGTCTCTTTCATCTG 48899
Qy 205 sMetValPheGlyMetSerLeuSerProSerProAspTyrArgIleMetLeuGlyValle 225
Db 48900 CATGGTATTTGGCATGTCCCTCATGCCACACGAGATGGAGGATCATGCTTGGGTTCT 48959
Qy 225 uAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuProGlnLeuSerProAr 245
Db 48960 ATCAATACCATCATATATATCTTGTGATTCACCATCTCTTTTACTTACCTGTAATCCGCGAG 49019
Qy 245 gTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGl 265
Db 49020 GTGGCTCGTGAGCAAGAAATGGCTGAGGCCACAGCGTGTGTGCAAGCGCTTCGCTGG 49079
Qy 265 yLysAspValSer----- 270
Db 49080 AAGAGAAGATGTTTCAGGTTTGTATGTCTTACCACGACATTTTCCCTGTGAGGATCTAGA 49139
Qy 270 ----- 270
Db 49140 AAGTGGAGAAACTATTTAATGCTATCAGTATAGGTATATATCTATGCGATGGAGTTTCT 49199
Qy 271 -----GlyGluLeuSerLeuLeuGluGlyLeuGluValGlyG 284
Db 49200 CATGTTGAATTAACATGCTCCAGGAAATGGCCCTTCTCGTTGAAGGTCTGGGGTTGGCA 49259
Qy 284 yAspThrSerIleGluTyrIleLeuGlyProAlaThrGluAlaAlaAspAspLeuV 304
Db 49260 AAGACACAAAAATTGAGGAATACATAATTGGACCTGATGATGAGCTTCTCTGATGAAGGCG 49319
Qy 304 aThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpI 324
Db 49320 TGGCTCCA---GATCCAGAGAGATCAACTGTATGTCCTGAAGAGCTTATCGTGG 49376
Qy 324 leAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgH 344
Db 49377 TTGCCCGTCTGTTCACGGCAAGTGCATCTGGAAGTGCATTAGTCTCATCTCTCTC 49436
Qy 344 isGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlyS 364
Db 49437 ATGGTAGTATGTCATGCTCAGGTAAGCCCTTGTGGATCTCTGTGTCTACCCCTTTTGA 49496
Qy 364 erValHisGluAsnMetProGlnAlaGlyLysSerMetArgSerThrLeuPheProAsnP 384
Db 49497 GTGTCCATGAGAAGATGCTTGAGATAATGGAAGCATCGGAGCACATTTGTTCTTA 49556
Qy 384 heGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluA 404
Db 49557 TTGGCAGCATGTTTATGTGCGGAACAGCAGCAAGCTTAAAGGTGATTGGGATGCTGAGA 49616
Qy 404 snLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspA 424
Db 49617 GT---CAACGGAGGGGTGAAGTATGATCAGACCATGCTGGGATCAGATTGAAGATA 49673
Qy 424 snLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIle---V 443
Db 49674 GCCTCCAAAGGCCCACTTATTTCTGCTCAAGCGACAAAGCGTGGAAAGAGATCGCTG 49733
Qy 443 aHisHisGlyHisArgGlySerAlaLeuSerMetArgGlnThrLeuLeuGlyGluG 463
Db 49734 CACCTCATGGCAGTATATGGTGTCTGTGGGA-----AGAAGTAGTAGTCTCATCGCAGG 49787
Qy 463 yGlyAspGlyValSerSerThrAspIleGlyGlyTyrPdnLeuAlaTrpLysTrps 483
Db 49788 GCGGGGAGGACAGTAAAGCAGCATGGCAATTGGTGGGGATGGCAGTTGGCTTGGAAATGGA 49847
Qy 483 erGluLysGluGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisG 503

Db 49848 CTGAGCAGAGGTCAGATGCGGAAAAAGAGGTCGCTTCCAAAGTACTACTTCCATG 49907
 Qy 503 lngluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyAspV 523
 Db 49908 AAGAGGCTGTGACAGGTGATCGCAGGGGCTCTATCTGTCATTGCTCT--GGAGGTGATG 49964
 Qy 523 alPheGluGlySerGluPheValHisAlaAlaValSerGlnSerAlaLeuPheS 543
 Db 49965 TTCTCTCTGTGTGTGAGTTCGTCCAGCAGCTGCTCTGTGACCAACCTGCTCTTTACT 50024
 Qy 543 exlYsGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValA 563
 Db 50025 CTAAGGAATTGATGGACCAAGCCTTGCTGGCCCTGCTATGTTGTCATCTCAGGCAG 50084
 Qy 563 laAlaIysGlySerArgTTPlyAspLeuPheGluProGlyValArgAlaLeuLeuV 583
 Db 50085 TTGCTAAAGGTCAAAATGGCGACACTTATTCGAACCTGGAGTGAAGCATGCTCTGTTG 50144
 Qy 583 alGlyValGlyVileGlnIleLeuGln----- 591
 Db 50145 TTGGCATTAGGATACAAATCCCGACAGTAACCCCTTTTCTGATTTTATGTCATGTAC 50204
 Qy 591 ----- 591
 Db 50205 GAATGTCATGTGCTATGTTGAACCTAAATTTTCTAGTAGCTCCAGTATAAATTAATCTGAT 50264
 Qy 591 ----- 591
 Db 50265 TAATGTAACCTGGTTAAATCAAGCATCTTCATCGAGTGTCTCTAGTCTTCTAGCGAG 50324
 Qy 591 ----- 591
 Db 50325 TCTCTCAATAAGCTCTTACTTACTTACTCGAAAAACATGTTTAGCTGTTAGTGTGTTT 50384
 Qy 591 ----- 591
 Db 50385 TCTTGCATATATATTCATACAAAGCAATGAGCTAAACCTGCACAAATAAATACACTTTG 50444
 Qy 591 ----- 591
 Db 50445 ATGATTAGATACAAATGCCATCAATGTTGTTAGTCTTTTACACTAGTAAAGTCAACT 50504
 Qy 591 ----- 591
 Db 50505 TAACATCTTGCAGTATTCATATTAAGAAACCAATGGGACTGGGCTATGTCACATGACC 50564
 Qy 591 ----- 591
 Db 50565 TGATTACGGGTATCTATCTGAGCAGACAGAAACAATGTGCATGCTCTTTTATTCCTG 50624
 Qy 591 ----- 591
 Db 50625 AATTTATTTGCTGAATATGTTCTAGTTTCGAGAAATAACTTTTCAGAAACTTTTCACAGT 50684
 Qy 591 ----- 591
 Db 50685 GGCATAATCATATATCTCTTTTACTTCCCTCCACTCTCTCTATTTTCCCTTGTAT 50744
 Qy 591 ----- 591
 Db 50745 ATTTTATATCAGTCTATTCATGTTTACTTACTTACCTCACTATAGCATACTGAATGTTAATGTCAT 50804
 Qy 592 -----GlnPheAlaGlyIleAsnGlyValLeuTyrrThrProGln 605
 Db 50805 CAACCTTTCCATATGTCGAGTGTGTCGATTAATGAGTCTCTGTACTACACTCCACAA 50864
 Qy 606 IleLeuGluGlnAlaGlyValAlaValIleLeuSerIysPheGlyLeuSerAlaSer 625
 Db 50865 ATTTCTTGAGCAAGCTGGTGTGTTGTTCTTCTTTGCAACATTTAGGACTTCTCTCATCT 50924
 Qy 626 AlaSerIleLeuSerSerLeuThrLeuLeuMetLeuProCysIleGlyPheAla 645
 Db 50925 GCATCTATTCTTATTAGCGGACTGACAACTTGTGCTGATGCTTCCCGACGATTCGTTATGCT 50984

Qy 646 MetLeuLeuMetAspLeuSerGlyArg----- 655
 Db 50985 ATGAGGCTCATGGATATGTCCTGGAAGAAGGTAGACTTCATATTTTACATTTTTTTTGT 51044
 Qy 655 ----- 655
 Db 51045 GATTACATGCATCACTTCTAAGTCGAAAGTATTATGCTAACCATTAATTCCTGGTGTCTTCT 51104
 Qy 656 -----PheLeuLeuLeuGlyThrIleProIleLeuLeuAlaSerLeuValIleLeu 672
 Db 51105 CATTCAGGTTTCTT-CTCTTTCGCAACAATCCCTATCCTGATAGTAGTACTAGTATCTTG 51163
 Qy 673 ValValSerAsnLeuLeuAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSer 692
 Db 51164 ATTCTGTCATATTCGGATGTCGGGACCATGTTTCATGCTCCTCAGTGTCCAGTCACT 51223
 Qy 693 ValIleValTyrrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys 712
 Db 51224 GTCATACTCTACTTCTCTTCTTGTCTGTCGGGTTTCGGGCTATTCCAAACATTTCTCTGT 51283
 Qy 713 AlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPhe 732
 Db 51284 GCAGAGATTTTCCGACACCCGTCGTCGTCATGTCATGTCATGTCGTCCTTAAACATTC 51343
 Qy 733 TriPleGlyAspIleIleValThrTyrrSerLeuProValMetLeuAsnAlaIleGlyLeu 752
 Db 51344 TGGATCGGTGATATCATTTGTGACATACACCTCCCGTGATGCTCAACGCCATTGGACTC 51403
 Qy 753 AlaGlyValPheSerIleTyrrAlaValValCysLeuLeuSerPheValPheValPheLeu 772
 Db 51404 GCTGAGTGTGTTGAAATCTACGAGTGTGTCATATGCTTTCCTGTTTCTTCTCATG 51463
 Qy 773 LysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGly 792
 Db 51464 AAGTGGCGGAGACAAAGGCGATGCTTCTTGAAGTCAATCCAGATCTTCTCTGTCGGA 51523
 Qy 793 AlalysGlnAla 796
 Db 51524 GCAAAGCAGGCC 51535
 RESULT 15
 AUF23E12 86710 bp DNA linear PLN 01-APR-1999
 LOCUS Arabidopsis thaliana DNA chromosome 4, BAC clone F23E12 (ESSA project).
 DEFINITION AL022604
 ACCESSION AL022604.1 GI:3080406
 VERSION AL022604
 KEYWORDS Arabidopsis thaliana (thale cress)
 SOURCE Arabidopsis thaliana
 ORGANISM Arabidopsis thaliana
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 REFERENCE 1 Bevan, M., Hilbert, H., Braun, M., Holzer, E., Brandt, A., Duesterhoeft, A., Hohnel, J., Mewes, H.W., Mayer, K.F.X. and Schueller, C.
 AUTHORS EU Arabidopsis sequencing, project.
 JOURNAL Submitted (01-APR-1999) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG. E-mail: schueller@mips.biochem.mpg.de, mayermips.biochem.mpg.de
 REFERENCE 2 (bases 1 to 86710)
 AUTHORS Direct Submission
 TITLE EU Arabidopsis sequencing, project.
 JOURNAL
 COMMENT E-mail: michael.bevan@bbsrc.ac.uk
 Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>.
 FEATURES Location/Qualifiers
 source 1..86710


```
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variant="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
173..1890
/gene="F23E12.10"
join(173..374,456..919,1089..1610,1678..1890)
/gene="F23E12.10"
/codon_start=1
/product="hypothetical protein (fragment)"
/protein_id="CAA18726.1"
/db_xref="GI:3080407"
/db_xref="GOA:O65486"
/db_xref="TrEMBL:O65486"
/translation="NAAVAGCFEAVESVLPSSSTDSSTLPNTTSMVILSAVTASV
VSEILGSPAPKVPDPSPGELPLVLLGALGLVSLALSRCTSSMTSAVDLSLNK
DAGIAPKAVFPVGGSLVGIILVYEVLYWGFQNDVILLEKRPVKGLSADLLQLVA
VIAATAWCRASGLVGGYVAPSLFVGGAGMAYGKFIPLGALQCNDFNLISLEVASPQ
AYGLGMATLAGVQVPLTAVLLFELTQDYRIYVPLLGAVGMSWITSGSKRQET
RRTKTRKKSQEAQVSTSDDESSTNNCEVESLCLDDSLNQSEELPKSIFVSEA
MRTFRATVMSTSLREALTRMLIEKQSCALIVDPNIFILITLSDIOEFKARKEGN
NPKDIFVNDICSRSGKCKVPWTVPDMLLAAQTINMKHLSHVAVVSGSIDAPRI
HVGVLDRSCILTRR"
173..374
/gene="F23E12.10"
/number=1
375..455
/gene="F23E12.10"
/number=1
456..919
/gene="F23E12.10"
/number=2
920..1088
/gene="F23E12.10"
/number=2
1089..1610
/gene="F23E12.10"
/number=3
1611..1677
/gene="F23E12.10"
/number=3
1678..1890
/gene="F23E12.10"
/number=4
3124..4428
/gene="F23E12.20"
join(3124..3241,3323..3465,3595..3696,3945..4107,
4211..4428)
/gene="F23E12.20"
/number=1
/note="similarity to dihydroflavonol-4-reductase, Gerbera
x sp., F1R2:S35189"
/codon_start=1
/product="putative protein"
/protein_id="CAA18727.1"
/db_xref="GI:3080408"
/db_xref="TrEMBL:O65487"
/translation="MQQAKGVCTGATGASGLASWLVKRLILLEGYEVIGTVRDPGNEKK
LAHLNKLGEAKRLVADLMEEGSDFNAMCCQGFHTASPIRLPAIGETINVLK
SCRNPKLKEVLTSSSTVWYALSKTLAEQAAKFESENGIDLTVLFLFVGLSLP
PDLCSATSDYLGILKGETEKFGWQGGYVHDDVARTHIVVFEHAAQGRYCISNV
ISLEELVSLFASRYPSLIPIKRFDTPLN"
3124..3241
/gene="F23E12.20"
/number=1
3242..3322
/gene="F23E12.20"
/number=1
3323..3465
/gene="F23E12.20"
/number=2
3466..3594
/gene="F23E12.20"
/number=2
```

```
/number=2
3595..3696
/gene="F23E12.20"
/number=3
3697..3944
/gene="F23E12.20"
/number=3
3945..4107
/gene="F23E12.20"
/number=4
4108..4210
/gene="F23E12.20"
/number=4
4211..4428
/gene="F23E12.20"
/number=5
4939..6163
/gene="F23E12.30"
complement(join(4939..5013,5109..5195,5307..5342,
5451..5651,5725..5811,6161..6163))
/gene="F23E12.30"
complement(join(4939..5013,5109..5195,5307..5342,
5451..5651,5725..5811,6161..6163))
/gene="F23E12.30"
/note="Contains Clathrin adaptor complexes small chain
signature [VYKRYASLYF]
contains EST gb:T44219"
/codon_start=1
/product="clathrin assembly protein AP19 homolog"
/protein_id="CAA18728.1"
/db_xref="GI:3080409"
/db_xref="GOA:O23685"
/db_xref="TrEMBL:O23685"
/translation="MIHFVLLVSRQGVLTQKVPYAKERSKVIRELSGVLNRGP
KLNFVWRGVKVVYKRYASLYFCMCDQSDNELEVEIHHYVEILDYRGVSCELD
LIFHFKAAYILDELLIAGLQSSKKTVAIIISAQDLVEVAKEASSISNIQAQT
NR"
complement(4939..5013)
/gene="F23E12.30"
/number=1
complement(5014..5108)
/gene="F23E12.30"
/number=1
complement(5109..5195)
/gene="F23E12.30"
/number=2
complement(5196..5306)
/gene="F23E12.30"
/number=2
complement(5307..5342)
/gene="F23E12.30"
/number=3
complement(5343..5450)
/gene="F23E12.30"
/number=3
complement(5451..5651)
/gene="F23E12.30"
/number=4
complement(5652..5724)
/gene="F23E12.30"
/number=4
complement(5725..5811)
/gene="F23E12.30"
/number=5
complement(5812..6160)
/gene="F23E12.30"
/number=5
complement(6161..6163)
/gene="F23E12.30"
/number=6
7352..7582
/gene="F23E12.40"
complement(7352..7582)
```


QY 533 laAlaLeuValSerGlnSerAlaLeuPheSerIysGlyLeuAlaGluProArgMetSerA 553
 Db 41305 CTCTTTGGTTAGCCACGAGCTCTTTATTCARAGACCTTCTCAAGACATACAAATG 41364
 QY 553 spAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuP 573
 Db 41365 GTCTCTGATGTTACATCCATCCGAA--ACAACAAAGGGTCAATTTGGCATGATCTTC 41421
 QY 573 heGluProGlyValArgAlaLeuLeuValGlyValGlyLeuGlnIleLeuGln--- 591
 Db 41422 ATGATCTCGAGTCAAGCGTGCATTAGTCGTAGAGTTGAGCTTCAAACTACCTCAGCAG 41481
 QY 591 ----- 591
 Db 41482 TAGCTAGATAGCCACATCAAAATTTAGTGAATCACAGAGAAATATGCGTTTATGCTA 41541
 QY 592 -----GlnPheAlaGlyIleAsnGlyValLeuT 601
 Db 41542 ATTCTATGTTTCTCTCTCTCTGAAATGAGTTCTCAGGCATCAACGGAGTTCTTT 41601
 QY 601 YrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyL 621
 Db 41602 ACTACACCCGCAATCCCTTACGACGCGGGTTCGGATCTTACTATCGAACATGGGA 41661
 QY 621 euSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProc 641
 Db 41662 TTAGTTCTCTCAGCATCCTTACTTATAAGTCAATGACAACTTTGTGATGTTACCTG 41721
 QY 641 ysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg----- 654
 Db 41722 CAATAGCTGTGCAATGAGGCTCATGGATCTTTCTGTCGAGGCTGATTAACCTAAAT 41781
 QY 654 ----- 654
 Db 41782 TTCTTTTGTCTTACATATACATACATAACAGAAACAGCTAAATGCGTTATGAAATGCTT 41841
 QY 655 -----ArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeu 669
 Db 41842 TTATTTTATATTGTCAGGACCTTGTCTTCCACGATACCAATCCCTGATAGCATCTCTA 41901
 QY 670 ValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSer 689
 Db 41902 TTGGTTTATAGTAACTCAATCTTGTTCAATGACAGCATTTGTCACGCGGCTTATCA 41961
 QY 690 ThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsn 709
 Db 41962 ACCGTAAGCGTTGTCTACTTCTGCTTCTGCTGATGGGTTTCGGTCTCTCTCCAAAC 42021
 QY 710 IleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAla 729
 Db 42022 ATCTCTGTTCAGAGATTTTCCAACTCGAGTCCGGAACTGCAATCGCAATCTGGCA 42081
 QY 730 PheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAla 749
 Db 42082 CTCACCTTCGTGATCTGTGACATAATCGTCACTTACAGTCTCCCGGTGCTCAATCC 42141
 QY 750 IleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPhe 769
 Db 42142 ATTGACTAGCTGGTGTGTTTGGAAATGACCAATCGTATGTTGCATTTTCATGGGCTTTT 42201
 QY 770 ValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePhe 789
 Db 42202 GTGTTTCATTAAGTCCCGGAACCTAAGGCATGCCACTTGAAGTCATCACAGAGTTCTTT 42261
 QY 790 AlaValGlyAlaLysGlnAlaAla 798
 Db 42262 TCTGTGGAGCTAGACAGCTGAAGCT 42288

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 13, 2004, 11:39:19 ; Search time 978.566 Seconds
(without alignments)
4291.521 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSWLAVQTFPTPLDR.....PLEVITEFPAVCAKQAARA 800

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 segs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2/1/USPFO.spool/US10051909/runat_13102004_123335_19900/app.query.fasta_1.1678
-DB=N Geneseq 23Sep04 -START=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPT=0
-LOOPT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blcsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10051909 @CGN 1.1 891 @runat_13102004_123335_19900 -NCPU=6 -ICPU=3
-DEV TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 23Sep04:*

1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001bs:*
5: geneseqn2002bs:*
6: geneseqn2003bs:*
7: geneseqn2004bs:*
8: geneseqn2005bs:*
9: geneseqn2006bs:*
10: geneseqn2007bs:*
11: geneseqn2008bs:*
12: geneseqn2009bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4075	100.0	2777	12	ADG47935 Corn Arab
2	4059	99.6	2908	12	ADM47932 Polynucle
3	3517	86.3	2824	6	ABK51962 Corn cont
4	3517	86.3	2824	8	ABX93198 cdna enco
5	3517	86.3	2824	12	ADG47905 Corn Arab
6	2674	65.6	2601	6	ABK51965 Soybean c

7	2674	65.6	2601	8	ABX93201	ABX93201 cdna enco
8	2674	65.6	2601	12	ADG47911	Adg47911 Corn Arab
9	2584	63.4	2190	6	ABZ13510	Abz13510 Arabidops
10	2275	55.8	2205	6	ADG87978	Adg87978 A. thalia
11	2275	55.8	2205	6	ADG87978	Adg87978 A. thalia
12	1670.5	41.0	1692	8	ABK51966	Abk51966 Soybean c
13	1670.5	41.0	1692	8	ABX93202	ABX93202 cdna enco
14	1670.5	41.0	1692	12	ADG47913	Adg47913 Soybean A
15	1424	34.9	1487	8	ABK51968	Abk51968 Wheat con
16	1424	34.9	1487	8	ABX93204	ABX93204 cdna enco
17	1424	34.9	1487	12	ADG47917	Adg47917 Wheat Ara
18	1035	25.4	1412	12	ADJ39747	Adj39747 Plant CDN
19	943	23.1	1009	6	ABK51969	Abk51969 Wheat CDN
20	943	23.1	1009	8	ABX93205	ABX93205 cdna enco
21	943	23.1	1009	12	ADG47919	Adg47919 Wheat Ara
22	620	15.2	751	12	ADJ41683	Adj41683 Plant CDN
23	616	15.1	1806	10	ADC07791	Adc07791 Rice DNA
24	600	14.7	778	12	ADJ41684	Adj41684 Plant CDN
25	587.5	14.4	1518	6	ABL41880	AbL41880 Nucleotid
26	562	13.8	870	6	ABK51964	Abk51964 Rice CDNA
27	562	13.8	870	8	ABX93200	ABX93200 cdna enco
28	562	13.8	870	12	ADG47909	Adg47909 Rice Arab
29	552	13.5	659	12	ADJ41685	Adj41685 Plant CDN
30	548	13.4	644	12	ADJ42193	Adj42193 Plant CDN
31	539	13.2	1395	6	ABK73616	Abk73616 Bacillus
32	534	13.1	2127	12	ADM47936	Adm47936 Polynucle
33	520.5	12.8	1826	3	AAC42332	Aac42332 Arabidops
34	496	12.2	1853	6	ABK51973	Abk51973 Soybean c
35	496	12.2	1853	8	ABX93209	ABX93209 cdna enco
36	496	12.2	1853	12	ADG47927	Adg47927 Soybean B
37	494.5	12.1	2017	6	ABK51972	Abk51972 Rice cont
38	494.5	12.1	2017	8	ABX93208	ABX93208 cdna enco
39	494.5	12.1	2017	12	ADG47925	Adg47925 Rice Beta
40	493.5	12.1	1527	3	AAC43261	Aac43261 Arabidops
41	493	12.1	1482	12	ABZ12990	Abz12990 Arabidops
42	493	12.1	1482	12	ADN72404	Adn72404 Thale cre
43	492	12.1	1338	8	ACA47151	AcA47151 Prokaryot
44	491.5	12.1	1914	6	ABK51971	Abk51971 Corn CDNA
45	491.5	12.1	1914	8	ABX93207	ABX93207 cdna enco

ALIGNMENTS

RESULT 1

ADG47935
ID ADG47935 standard; cdna; 2777 BP.

XX
AC ADG47935;

DT 11-VAR-2004 (first entry)

XX Corn Arabidopsis-like sugar transport protein cdna #2.

XX Arabidopsis-like sugar transport protein;
KW Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;
KW corn; plant; gene; ss.

XX Zea mays.

XX Key Location/Qualifiers

FT CDS 12..2414

FT /tag= a

FT /product= "Corn Arabidopsis-like sugar transport protein"

FT /note= "No start codon"

FT /partial

XX US2002199217-A1.

XX 26-DEC-2002.

XX 17-JAN-2002; 2002US-00051903.

XX 24-APR-1998; 98US-0083044P.

PR 14-APR-1999; 99US-00291922.
 XX (HELE/) HELENTJARIS T G.
 XX Helentjaris TG;
 DR WPI; 2004-040967/04.
 DR P-PSDB; ADG47936.
 XX
 PT New isolated polynucleotide encoding a polypeptide having sugar transport
 PT protein activity, for producing a transformed plant and for use as probes
 PT in physical mapping.
 XX
 PS Claim 6; SEQ ID NO 31; 71pp; English.
 XX
 CC The invention relates to Arabidopsis-like or Beta-vulgaris-like sugar
 CC transport proteins and their corresponding nucleic acid sequences. The
 CC transgenes of the invention are useful to transform a cell. These are also
 CC useful to produce a transgenic plant. Probes derived from sequences
 CC encoding sugar transport protein may be used for physical mapping. The
 CC present sequence is corn Arabidopsis-like sugar transport protein cDNA.
 XX
 SQ Sequence 2777 BP; 667 A; 587 C; 701 G; 822 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0 Length: 2777
 Score: 4075.00 Matches: 800
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0
 US-10-051-909-32 (1-800) x ADG47935 (1-2777)
 QY 1 IleArgSerGlySerTrpLeuAlaValGlnThrProPheThrProAspLeuAspArg 20
 Db 12 ATTCGGAGCGCTCTGGCTTGCAGTCCAAACGCCCTTCACCCCTGATCGACCGGAGG 71
 QY 21 GluArgLeuLeuProSerValValLeuAlaLeuProGlyProLeuProPheThrProAspLeu 40
 Db 72 GAGCGGCTCTCCGTCAGTGTCTTCTTGTCTTGGCTTGGCGCTCTTCCGCTCGTCTCGTGT 131
 QY 41 SerSerGlnGluProValThrSerAspAlaLeuGluAspLysMetSerGlyAlaVal 60
 Db 132 TCTTCACAGAGCGGTGACCTCGACGATATCTTCGAGGACAGATGTCGGGGGCTGT 191
 QY 61 LeuValAlaValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIle 80
 Db 192 CTTGTGCGCCATAGTCGCTCCATCGGCAATCTATTGCGGGGTGGGACCAATGCCACCATC 251
 QY 81 AlaAlaValLeuTrpIleLysLysGluPheGlnLeuGlnAsnGluProThrValGlu 100
 Db 252 GCAGCTGCTGTTCTGTATATPAGAGAGATTTCAATTCGAAATGAGCCCACTGTGGAG 311
 QY 101 GlyLeuValSerMetSerLeuIleGlyAlaThrIleValThrPheSerGlyPro 120
 Db 312 GGACTAATTGTGTCAATGCTATTCGCGGCCACCATCGTTACTACATTCCTCGGGCCA 371
 QY 121 LeuSerAspSerIleGlyArgProMetLeuIleLeuSerSerIleLeuTrpPhePhe 140
 Db 372 TTATCAGATCGATTGGCGCGCCCTATGCTTATCTCTCTCAATCTGTACTTCTTC 431
 QY 141 SerGlyLeuIleMetLeuTrpSerProAsnValTrpValLeuLeuAlaArgPheVal 160
 Db 432 AGCGGCTCATCATGCTATGCTCTCTAATGCTATGCTCTGCTGTGGACGCTTCGTA 491
 QY 161 AspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrlleSerGluIleAla 180
 Db 492 GATGGATTGGATTGGCTGTGCTGACGCTGTGCTTGTGCTTGTACATTCAGAAATAGCC 551
 QY 181 ProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMet 200
 Db 552 CCTTCGAGATTAGAGTTTGTGTAATACACTACCAATTCAGTGGATCAGGAGGATG 611

QY 201 PheLeuSerTyrCysMetValPheClyMetSerLeuSerProSerProAspTrpArgIle 220
 Db 612 TTTCTTGTCTATCTGTCATGCTGTTTGGGATGTCCTGTGCGCATCACCGATTGAGAAAT 671
 QY 221 MetLeuGlyValLeuAlaIleProSerLeuPhePhePheGlyLeuThrIlePheTyrlleu 240
 Db 672 ATGCTTGGTGTCTGCGCATACCTTCATTGCTTCTTGTGTTGACATATTTATCTT 731
 QY 241 ProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValleu 260
 Db 732 COTGAATCTCCAAGATGCTGCTTAGCAAGCTGGATGGCAGAGGCAAAAAGGTGTG 791
 QY 261 GlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeuLeuGluLysLeu 280
 Db 792 CAAAGTTTACGGGGGAAAGACGATGCTCAGTGTGATTTGCTCTTCTCGAAGGGTTG 851
 QY 281 GluValGlyGlyAspThrSerIleGluGlyTrpIleLeuGlyProAlaThrGluAlaAla 300
 Db 852 GAGGTGGAGGAGACACTTCCATTGAAGAGTACATCATTTGGACCTGCCACCGAGGAGCC 911
 QY 301 AspAspLeuValThrAspGlyAspLysGluGlnIleThrIleTyrlleuTyrlleuGly 320
 Db 912 GATGATCTTGTACTGACGCTGATAGGAAACAAATCACCTTTATGGGCTTGAAGAAGGC 971
 QY 321 GlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeu 340
 Db 972 CAGTCATGGATTGCTCGACCTTCTAAGGACCCCATCATGCTTGAAGTGTGCTTCTCTT 1031
 QY 341 AlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThr 360
 Db 1032 GCATCTCGTCATGGGAGCATGTTGAACAGAGTACCCCTTATGGATCCGATGTGTACA 1091
 QY 361 LeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeu 380
 Db 1092 CTTTITGGTAGTCTCATGAGATATCCCTCAAGCTGGAGAGATATGAGGACACATTG 1151
 QY 381 PheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluInTrp 400
 Db 1152 TTTCCAAACTTGGAGTATGTTTCAGTGTACAGATCAGCATGCCAAAATGAGCAGTGG 1211
 QY 401 AspGluGlnAsnLeuHisArgAspAspGluGlyTyrlleuAlaSerAspGlyAlaGlyAsp 420
 Db 1212 GATGAAGAGATCTTCATAGGGATGACAGAGTACCGATCTCATGTTGTCAGAGGTGAC 1271
 QY 421 TyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLys 440
 Db 1272 TATGAGACAAATCTCCATAGCCCATTTGCTCCAGGCGAGGCAACAGTGGGAGGGAAG 1331
 QY 441 AspIleValHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu 460
 Db 1332 GACATTTGACCATGCTGCTACCGTGGAGTGTCTTGAGCATGAGAGGCAAAACCTCTTA 1391
 QY 461 GlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyTrpGlnLeuAlaTrp 480
 Db 1392 GGGAGGCTGCAGATGCTGTGAGCAGCACTGATATCGTGGGGGATGGCAGCTTGTGG 1451
 QY 481 LysTrpSerGlnLysGluGlyGluAsnGlyArgLysGluGlyClyPheLysArgValTyrl 500
 Db 1452 AAATGGTCAGAGAGGAGGTGAGATGGTAGAAGGAGGAGGTTTCAAAAGAGGTCTAC 1511
 QY 501 LeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly 520
 Db 1512 TTSACCAACAGAGGAGTCTCTGCTCAAGAGGGGCTCAATTTGTTTCACTTCCCGTGT 1571
 QY 521 GlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAla 540
 Db 1572 GGCAGTGTCTTTGAGGGGTAGTGTGATTTGATCATGCTGCTCTTTAGTAGTCAGTCAGCA 1631
 QY 541 LeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSer 560
 Db 1632 CTTTITCTCAAAGGGTCTTGTCTGAACCAACGATGTCAGATGCTGCCATGTTTCAACCATCT 1691

QY 561 GluValAlaAlaLysGlySerArgTTPLeuPheGluProGlyValArgAla 580
DB 1692 GAGTAGCTGCCAAGGTTACGTTGGAAGATTGTTTGACCTGGAGTGGCGTGCC 1751
QY 581 LeuLeuValGlyValGlyLeuGlnLeuGlnPheAlaGlyLeuGlnValLeu 600
DB 1752 CTGTTAGTCGGTCTGGAATTACAGATCTTACAGATTTGCTGGATAAACGGTGTCTG 1811
QY 601 TTTTyrThrProGlnLeuGlnAlaGlyValAlaValLeuSerLysPheGly 620
DB 1812 TACTATACCCCAAAATCTTGAGCAAGCTGGTGGAGATTATCTTTCCAAATTGGT 1871
QY 621 LeuSerSerAlaSerAlaSerLeuLeuLeuSerSerLeuThrThrLeuLeuMetLeuPro 640
DB 1872 CTCAGCTCGGCATCAGCATCCATCTTGATCAGTTCTCTCACTACCTTACTAATGCTTCT 1931
QY 641 CysIleGlyPheAlaMetLeuMetAspLeuSerGlyArgArgPheLeuLeuGly 660
DB 1932 TGCATTGGCTTTGCCATGCTGATGAGATCTTCCGAGAGAGGTTTTCGCTGAGGC 1991
QY 661 ThrIleProIleLeuLeuAlaSerLeuValIleLeuValValSerAsnLeuLeuAspLeu 680
DB 1992 ACAATCCATCTTGTAGCATCTCTAGTTATCTCTGTTGTGTCATCTAATTTGATTG 2051
QY 681 GlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPhe 700
DB 2052 GGTACACTAGCCCATGCTTTGCTCTCCACCGTCAGTGTATCTCTCTCTCTCTCTCT 2111
QY 701 ValMetGlyPheGlyProIleProAsnLeuLeuCysAlaGluIlePheProThrArgVal 720
DB 2112 GTTATGGGATTTGGTCCATCCCAATTTTATGTGAGATCTTTCACACAGGGTT 2171
QY 721 ArgGlyLeuCysLeuAlaIleCysAlaPheThrPheThrPheIleGlyAspIleLeuValThr 740
DB 2172 CGTGGCCTCTGTATTGCCATTGTGCTTTACATCTCTGGATCGAGATATCATCTGCACG 2231
QY 741 TyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAla 760
DB 2232 TACAGCCTTCTGTGATGCTGGAATGCTATTGGACTGGCGGGTGTTCACATATATGCA 2291
QY 761 ValValCysLeuIleSerPheValPheLeuLysValProGluThrLysGlyMet 780
DB 2292 GTCGTAGCTTGATTTCTTTGTTGTTCTTCTTAAAGTCCCTGAGACAAAGGGGATG 2351
QY 781 ProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLysAla 800
DB 2352 CCCCTTGAGGTTATTACCGAATCTTTGCAATTTGTCGAGTGGTGGCAAGCGGCTGCAAAAGCC 2411

RESULT 2
ADM47932
ID ADM47932 standard; DNA; 2908 BP.
XX AC ADM47932;
XX DT 03-JUN-2004 (first entry)
XX DE Polynucleotide sequence #350 useful in producing transgenic plants.
XX KW Plant; maize; rice; sorghum; thale cress; soybean; wheat; transgenic;
XX KW osmotic stress; sugar transport; cell cycle pathway; plant height;
XX KW carbohydrate transport; crop productivity; plant growth;
XX KW stress resistance; disease resistance; insect resistance; heat tolerance;
XX KW nitrogen assimilation; water stress tolerance;
XX KW photosynthetic carbon fixation; virus resistance; gene therapy; gene; ds.
XX OS Zea mays.
XX FN US2003233670-A1.
XX PD 18-DEC-2003.
XX PF 04-DEC-2002; 2002US-00310154.
XX XX

04-DEC-2001; 2001US-0337358P.
(EDGE/) EDGERTON M D.
(CHOM/) CHOMET P S.
(LACC/) LACCETTI L B.
XX Edgerton MD, Chomet PS, Laccetti LB;
XX WPI; 2004-061374/06.
XX P-PSDB; ADM48300.
XX New polynucleotide, useful for manipulating plant protein quality,
XX improving plant growth, yield and crop productivity or grain composition
XX or producing plants with improved properties.
XX Claim 1; SEQ ID NO 350; 144pp; English.
XX The present invention relates to polynucleotide sequences, and the
XX proteins they encode. The sequences are isolated from a variety of
XX organisms such as plants (e.g. maize, rice, sorghum, thale cress,
XX soybean, and wheat), cyanobacteria, bacteria, yeast and other fungi. The
XX polynucleotide and polypeptide sequences of the invention are useful in
XX the production of transgenic plants that have improved properties. Also
XX disclosed are methods of producing fertile transgenic plants, preferably
XX maize, with desired phenotypes. The polynucleotide and polypeptide
XX sequences are useful for improving plants by providing protection against
XX osmotic stress, improving altering sugar transport and/or metabolism,
XX modifying the cell cycle pathway, reducing plant height, modifying plant
XX carbohydrate transport, improving crop productivity, improving plant
XX growth and stress resistance, improving disease resistance, improving
XX insect resistance, improving cold or heat tolerance, improving nitrogen
XX assimilation, improving stalk strength, improving water stress tolerance,
XX improving photosynthetic carbon fixation, improving biotic and abiotic
XX stress resistance, improving resistance to oxidative stress, providing
XX increased vigour, reducing senescence, and conferring virus resistance.
XX The present sequence represents a polynucleotide sequence of the
XX invention. Note: The sequence data for this patent is not provided in the
XX printed specification but is obtained in electronic format from the USPTO
XX website at seqdata.uspto.gov.
XX SQ Sequence 2908 BP; 669 A; 640 C; 747 G; 852 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 0 Length: 2908
Score: 4059.00 Matches: 797
Percent Similarity: 99.75% Conservative: 1
Best Local Similarity: 99.62% Mismatches: 2
Query Match: 99.61% Indels: 0
DB: 12 Gaps: 0
US-10-051-909-32 (1-800) x ADM47932 (1-2908)
QY 1 IleArgSerGlySerTyrLeuAlaValGlnThrProPheThrProAspLeuAspArg 20
DB 166 ATTCGAGCGGCTCTTGCTTCAGTCCAGAGCGCTTCACCCCTGATCTGGACCGGAGG 225
QY 21 GluArgLeuLeuProSerValValLeuAlaLeuProGlyProLeuProProAlaSerCys 40
DB 226 GAGCGCTCTTCCGTCAGTTGTTCTTGTTCCTGGGCGCTCTTCGCGCTCTCTGCTGT 285
QY 41 SerSerGlnGluProValThrSerAspAspIleLeuGluAspLysMetSerGlyAlaVal 60
DB 286 TCTTACAGAGCGGCTGACCTCGGACGATATCTTGGAGACACAGATGTCGGGGGCTGT 345
QY 61 LeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyrAspAsnAlaThrIle 80
DB 346 CTGTGCGCCATAGTCGCTCCATCGCAATCTATTGAGGGGTGGACAATGCCACCATC 405
QY 81 AlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGluProThrValGlu 100
DB 406 GCAGCTGCTGTTCTGTATATAAGAGGAATTCATATGCAAAATGAGCCCATCTGGAG 465
QY 101 GlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValIleThrPheSerGlyPro 120

```

Db 466 GGACTAAATGTGTCAATGTCACTTATCGGCGCCACCATCGTTACTATCACTCTCCGGGCCA 525
Qy 121 LeuSerAspSerIleGlyArgProMetLeuIleLeuSerSerIleLeuTyrPhePhe 140
Db 526 TTATCAGACTCGATTGGCGGCGCCCTATGCTTATCTCTCTCAATCTGTTACTTCTTC 585
Qy 141 SerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeuLeuAlaArgPheVal 160
Db 586 AGCGGCTCATCATGCTATGCTCTCTAAATGCTATGCTCTGCTGTGGCAGCGCTTCGTA 645
Qy 161 AspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAla 180
Db 646 GATGGAATTTGATATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 705
Qy 181 ProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMet 200
Db 706 CTTTCGGAGATTAGAGGTTTGTCTGAATACACTACCAATTCAGTGGATCAGGAGGAATG 765
Qy 201 PheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSerProAspTyrArgIle 220
Db 766 TTCCTGTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 825
Qy 221 MetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeu 240
Db 826 ATGCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 885
Qy 241 ProGluSerProArgTyrPheLeuValSerGlyArgMetAlaGluAlaLysValLeu 260
Db 886 CCTGAATCTCCAGATGGCTCGTTAGCAAAAGGTCGATGCGCAGAGGCAAAAGGTCGTTG 945
Qy 261 GlnLysLeuArgGlyLysAspValSerGlyLeuLeuSerLeuLeuGlyLeu 280
Db 946 CAAAGTTACGGGGAAAGAGATGCTCAGGTGAATTTGCTTCTTCGAGAGGTTG 1005
Qy 281 GluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAla 300
Db 1006 GAGTTTCGAGGAGACATCTTCATTTGAAGAGTACATCATTTGACCTGCCAGGAGGAGCC 1065
Qy 301 AspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGly 320
Db 1066 GATGATCTTGTACTGACGTGATAGAGAAACAATCACATTTATGGGCTGAGAGAGGC 1125
Qy 321 GlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeu 340
Db 1126 CAGTCATGGAATGCTCGACCTTCCAGGACCCAGCATGCTTGGAAAGTGTGCTTCTCTT 1185
Qy 341 AlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThr 360
Db 1186 GCATCTGCTCATGGAGCATGGTGAACCGAGAGTATCCCTTATGGATCCGATTGTGACA 1245
Qy 361 LeuPheGlySerValHisGluAsnMetProGlnAlaGlyLysSerMetArgSerThrLeu 380
Db 1246 CTTTTTGTAGTGTCCATGAGATATGCCCTCAGCTGGAGGAGTATGAGAGCATATG 1305
Qy 381 PheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrp 400
Db 1306 TTTTCCAAACCTTTGGAAGTATGTTTCAGTGTACAGATCAGCATGCGCAAAATGAGCAGTGG 1365
Qy 401 AspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAsp 420
Db 1366 GATCAAGAGAAATCTTATAGGATGACGAGGAGTACGCAATCTGATGCTCAGAGGTGAC 1425
Qy 421 TyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLys 440
Db 1426 TATCAGGACAAATCTCCATAGCCCATTTGCTGTCCAGGACGCAACAGGTGCGGAAGGGAAG 1485
Qy 441 AspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgGlnThrLeuLeu 460
Db 1486 GACATTTGTCACCATGCTACCGTGAAGTGTCTTTGAGCATGAGAGGCAAGGCTCTTTA 1545
Qy 461 GlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyTyrProGlnLeuAlaTrp 480

```

```

Db 1546 GGGAGGGTGGAGATGGTGTGAGCAGCACTGATATCGTGGGGGATGGCAGCTTGTCTGG 1605
Qy 481 LysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyr 500
Db 1606 AAATGGTCAGAGAAGAGGTGAGATGTTAGAAAGAGAGGTGGTTTCAAAGAGTCTAC 1665
Qy 501 LeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly 520
Db 1666 TTGCACCAAGAGGGAGTTCCTGGCTCAAGAAAGGGCTCAATTTGTTTCACTTCCCGGTGT 1725
Qy 521 GlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAla 540
Db 1726 GCGGATGTTCTTGGGGTAGTGAGTTTGTACATGCTGCTGCTTTAGTAGTCACTCAGCA 1785
Qy 541 LeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSer 560
Db 1786 CTTTTTCTCAAAGGGTCTTGTCTGAACCAACGATGTCAGATGCTGCATGTTCCACCATCT 1845
Qy 561 GluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAla 580
Db 1846 GAGTAGTCTGCCAAGGTTCAGTTGGAAAGATTGTTTGAACCTGGAGTGAGCGTGCC 1905
Qy 581 LeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeu 600
Db 1906 CTGTTAGTCTGGTGTGGAATTCAGATCCTTCAACAGATTGCTGGAATAAACGGTGTCTG 1965
Qy 601 TyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGly 620
Db 1966 TACTATATCCCAAAATCTTTCAGCAAGCTGGTGGCAGTATTTCTTCCAAATTTGGT 2025
Qy 621 LeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuPro 640
Db 2026 CTCAGTCTGGCATCAGCATCCATCTTGATCAGTCTCTCTACTACTTCTAATGCTTCT 2085
Qy 641 CysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGly 660
Db 2086 TGCATTGGCTTTGCCATGCTGCTTATGATCTTTCGGAAGAGGTTTTCGCTGCTAGGC 2145
Qy 661 ThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeu 680
Db 2146 ACAATTCATCTTGAATAGCATCTCTAGTATCTGTTGTTGTCCTCAATCTAATGATTG 2205
Qy 681 GlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPhe 700
Db 2206 GGTACACTAGCCCATGCTTGTCTCTCCACCGTCAGTGTATCGTCTACTTCTGCTGCTTC 2265
Qy 701 ValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgVal 720
Db 2266 GTTATGGGATTTGGTCCCATCCCAATTTATGTGTCAGAGATCTTTTCCACAGGGTT 2325
Qy 721 ArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThr 740
Db 2326 CGTGGCTCTGATTTGCCATTTGTGCTTTTACATCTCTGGATCGGAGATATCATCGTACC 2385
Qy 741 TyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAla 760
Db 2386 TACAGCTTCTCTGATGCTGAATGCTTATGGATGCTGCGGGTGTGTTTTCAGCATATATGCA 2445
Qy 761 ValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMet 780
Db 2446 GTGCTATGCTTGTATTTCTTTGTTGTTGCTTCTTCTTANGGTCCTGAGACAAAGGGGATG 2505
Qy 781 ProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLysAla 800
Db 2506 CCCCTTGAGGTATTTACCGAATCTTTGCAAGTGTGTCGGAAGCAAGCGGTGCAAAAGCC 2565

```

RESULT 3

ABK51962

ID ABK51962 standard; cDNA; 2824 BP.

XX ABK51962;

AC

DT 27-AUG-2002 (first entry)

XX Corn contig encoding Arabidopsis thaliana-like sugar transport protein.
DE Corn; Arabidopsis thaliana-like sugar transport protein;
KW carbohydrate transport; grain filling; annual field crop; plant; gene;
KW ss.
XX Zea mays.
OS
XX Key Location/Qualifiers
FH 238..2481
FT /*tag= a
FT /product= "Corn Arabidopsis thaliana-like sugar transport
FT protein"
XX US6383776-B1.
PN
XX 07-MAY-2002.
PD
XX 14-APR-1999; 99US-00291922.
PF
XX 24-APR-1998; 98US-0083044P.
PR
XX (DUPO) DU PONT DE NEMOURS & CO E I.
PA
XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;
PI WPI; 2002-453364/48.
DR P-PSDB; AAU97201.
DR
XX New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution.
PT
XX Claim 3; Col 25-28; 54pp; English.
PS
XX The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport protein
CC or Beta vulgaris-like sugar transport protein. The polynucleotide
CC sequences are useful for altering the level of sugar transport proteins
CC in plants, i.e. for control of carbohydrate transport and distribution in
CC plant cells, e.g. during grain filling of annual field crops (e.g. corn,
CC rice, soybeans, and wheat), and, for studying carbohydrate flows and
CC sugar transport. The polynucleotide sequences can also be used to isolate
CC cDNA sequences and genes that encode homologues of the new proteins. The
CC present sequence representing a contig assembled from various corn cDNA
CC clones encodes an Arabidopsis thaliana-like sugar transport protein
XX
SQ Sequence 2824 BP; 644 A; 649 C; 745 G; 775 T; 0 U; 11 Other;

Alignment Scores:
Pred. No.: 4,19e-298 Length: 2824
Score: 3517.00 Matches: 684
Percent Similarity: 95.18% Conservative: 27
Best Local Similarity: 91.57% Mismatches: 34
Query Match: 86.31% Indels: 2
DB: 6 Gaps: 2

US-10-051-909-32 (1-800) x ABK51962 (1-2824)

418 ACATCTCCGGGCGCCAAAGGGCTGACTCGTGGTAGGAGGCCCATGCTGGTCCGCTCGGCT 477
136 lleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu 155
478 GTCCTCTACTTCTGTCAGTGGGCTGTGTGATGCTTTTGGGCGCCAAATGTGTACATCTTGCTC 537
156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
538 CTCGCAAGGCTCATTTGATGGGTTCGGTATCGTTTGGCGGTACACACTTGTTCCTCTCTAC 597
176 lleSerGluIleAlaProSerGluIleArg--GlyLeuLeuAsnThrLeuProGlnPhe 194
598 ATCTCCGAAATGACACCGACAGANATCTTGGGGCTGNTNGAACACGCTTCCGCGAGTTC 657
195 SerGly---SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213
658 ATTGGGGTCAGNGGAGGGGATGTTCTCTCTCTACTGATGGTGTGGGATGTCCTCATG 717
214 ProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233
718 CCCAAACCTGATTGGAGGCTCATGCTGGAGTCTCTCGATCCGCTCACTTATTACTTT 777
234 GlyLeuThrIlePheTyrLeuProGluSerProArgTyrLeuValSerLysGlyArgMet 253
778 GGACTGACTGTCTTCTACTTGCCTGAATCACCAAGGTGGCTTGTNAGCAAGAGGAGATG 837
254 AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeu 273
838 GCGGAGGCGAAGAGAGTGTGCAAAAGGCTGCGGGGAGAGAGATGTCTCANGGAGANG 897
274 SerLeuLeuLeuGluGlyLeuValGlyGlyAspThrSerIleGluGlyTyrIleIle 293
898 GCTCTTCTAGTTGAAGGTTTGGGGTCGGTAAAGATACACGATTTNAGAGTACATCAAT 957
294 GlyProAlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGlnIleThr 313
958 GGACCTGCGCACCGAGGCGAGCGATGATCTTAACTGACGGTGATAGAGAAACAAATCACA 1017
314 LeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMet 333
1018 CTTTATGGCTGTGAAGAGGCGAGTCATGGATGCTCGACCTTCTAAGGGAGCCCATCATG 1077
334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
1078 CTTGGAAGTGTGCTTCTTCTTGCATCTGTCATGGAGCATGGTGAACACAGATGTACCC 1137
354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
1138 CTTATGGATCCGATTGTGACACTTTTGGTAGTGTCTCATGAGAATATGCTCAAGCTGGA 1197
374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
1198 GGAAGTATGAGAGACACATTTTCCAAACCTTTGGAAGTATGTTCACTGTACAGATCAG 1257
394 HisAlaLysAsnGluGlnTrpAspGluGlnAsnLeuHisArgAspAspGluGlyTyrAla 413
1258 CATGCCAAAATGAGCAGTGGGATGAAGAGAATCTTCATAGGGATGACGAGGATACGCA 1317
414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
1318 TCTGATGTGAGGAGGAGTACTATGAGACAATCTCCATAGCCCATTCCTGTCCAGGAG 1377
434 AlaThrGlyValAlaGluGlyLysAspIleValHisGlyHisArgGlySerAlaLeuSer 453
1378 GCAACAGGTGCGAAGGAGGAGACATTTGTGCACCATGTCACCGTGGAGTGTCTTGGC 1437
454 MetArgArgGlnThrLeuLeuGlyGlyAspGlyValSerSerThrAspIleGly 473
1438 ATGAGAGGCAAGGCTCTTATGGGAGGGTGGAGATGGTGTGAGCAGACACTGATATCGGT 1497
474 GlyGlyTyrGlnLeuAlaTrpIleSerGluLysGluGlyGluAsnGlyArgLysGlu 493
1498 GGGGGATGGCAGCTTGTGGAATGTCAGAGAGGAGGATGAGATGGTAGAAGGAA 1557

Qy 494 GlyGlyPheLysArgValTyrLeuHisGlnGluValProGlySerArgArgGlySer 513
 Db 1558 GGTGGTTTCAAAAGAGTCTACTTGCACCAAGAGGGAGTTCTTGGCTCAAGAGGGGCTCA 1617
 Qy 514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533
 Db 1618 ATTGTTTCACTTCCCGGTGGTGGCGATGTTCTTCAGGGGTAGTGAGTTGTATGCTGCT 1677
 Qy 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553
 Db 1678 GCTTTAGTAGTCAGTCAGCACTTTCTCAAGGGTCTTCTGCTGAACACGATGTCAGAT 1737
 Qy 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTPrLysAspLeuPhe 573
 Db 1738 GCTGCCATGTTTCAACCATCTGAGTAGCTGCCAAAGGTTACGTTGGAAAGATTGTTT 1797
 Qy 574 GluProGlyValArgArgAlaLeuValGlyValGlyValGlyLeuGlnLeuPhe 593
 Db 1798 GAACTGGAGTGAGCGTGGCTTGTAGTCGGTGTGGATTCAGATCCTTCAACAGTTT 1857
 Qy 594 AlaGlyLeuHisGlyValLeuTyrTyrThrProGlnLeuGluGlnAlaGlyValAla 613
 Db 1858 GCTGGATAAAGCGTGTCTGTACTATACCCCAAAATCTTGAGCAAGCTGGTGGCA 1917
 Qy 614 ValLeuSerLysPheGlyLeuSerSerAlaSerAlaSerLeuLeuSerSerLeu 633
 Db 1918 GTTATCTTTCCAAATTTGGTCTCAGCTGGCATCAGCATCCATCTGATCAGTCTCTC 1977
 Qy 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
 Db 1978 ACTACCTTACTATGCTTCTTGCATTTGGCTTGGCCATGCTGCTTATGATCTTTCCGA 2037
 Qy 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuLeuAlaSerLeuValIleLeuVal 673
 Db 2038 AGAAGGTGTTTGTGTGTAGCAAAATTCCTCAATCTTGATAGCATCTCTAGTTATCTGGT 2097
 Qy 674 ValSerAsnLeuLeuAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
 Db 2098 GTGTCCAATCTAATGATTTGGGTACACTAGCCATGCTTGTCTCCACCATCAGTGT 2157
 Qy 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnLeuLeuCysAla 713
 Db 2158 ATCGTCTACTTCTGCTCTCTGTTATGGGATTTGGTCCCATCCCAACATTTTATGTGCA 2217
 Qy 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTyr 733
 Db 2218 GAGATCTTTCCAAACGAGGTTCTGGCTCTGTATTTGCCATTTGCTTACATCTGG 2277
 Qy 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
 Db 2278 ATCGGAGATATCATCGTCACTACAGCTTCTGTGTATGCTGAATCTATTGACTGGCG 2337
 Qy 754 GlyValPheSerIleTyrAlaValValCysLeuLeuSerPheValPheValPheLeuLys 773
 Db 2338 GGTGTTTTCAGCATATATGAGTCGTATGCTTGAATTCCTTGTGTCTCTCTTATAG 2397
 Qy 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
 Db 2398 GTCCCTGAGACAAAGGGATGCCCTTGGATTATACCGAATTTCTTGCAGTTGGTGG 2457
 Qy 794 LysGlnAlaAlaLysAla 800
 Db 2458 AAGCAAGCGCTGCAAAAGCC 2478

RESULT 4

ABX93198
 ID ABX93198 standard; cDNA; 2824 BP.

XX
 AC ABX93198;

XX
 DT 29-MAY-2003 (first entry)

XX

DE CDNA encoding corn sugar transport protein #1.
 XX Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
 KW plant sugar transport protein; carbohydrate transport; soybean;
 KW carbohydrate distribution; plant; gene; ss.
 XX Zea mays.
 XX US2002178468-A1.
 XX 28-NOV-2002.
 XX 17-JAN-2002; 2002US-00051902.
 XX 24-APR-1998; 98US-0083044P.
 PR 14-APR-1999; 99US-00291922.
 XX (ALIE// ALLEN S M.
 PA (HITZ// HITZ W D.
 PA (KINN// KINNEY A J.
 PA (TING// TINGEY S V.
 XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;
 XX WPI; 2003-340957/32.
 DR P-PSDB; ABU08326.
 XX Novel plant sugar transport proteins and nucleic acid encoding the
 PT protein useful for producing transgenic plants having altered levels of
 PT sugar transport protein.
 XX Claim 2; Page 13-15; 56pp; English.
 XX The present invention relates to the isolation of Arabidopsis thaliana-
 CC like or Beta vulgaris-like sugar transport proteins, and the
 CC polynucleotide sequences encoding them. The plant sugar transport
 CC proteins of the invention have been isolated from corn, rice, soybean,
 CC and wheat. The polypeptides of the invention may be used for altering the
 CC level of expression of a sugar transport protein in a host cell, by
 CC transforming a host cell with a chimeric construct encoding all, or a
 CC portion of the sugar transport protein, in sense or antisense
 CC orientation. Particularly, the polypeptides may provide a means to
 CC control carbohydrate transport and distribution in plants. ABX93198-
 CC ABX93205 represent cDNA sequences encoding Arabidopsis thaliana-like
 CC sugar transport proteins
 XX Sequence 2824 BP; 644 A; 649 C; 745 G; 775 T; 0 U; 11 Other;

Alignment Scores:

Pred. No.: 4,19e-298 Length: 2824
 Score: 3517.00 Matches: 684
 Percent Similarity: 95.18% Conservative: 27
 Best Local Similarity: 91.57% Mismatches: 34
 Query Match: 86.31% Indels: 2
 DB: Gaps: 2

US-10-051-909-32 (1-800) x ABX93198 (1-2824)

Qy 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr 75
 Db 238 ATGGGGGCGCGGTAGTGTGCGCATCGCGCTCTATCGCAACTTGTGCGGGGTGG 297
 Qy 76 AspAsnAlaThrIleAlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
 Db 298 GACAAATGCAAAATGCTGAGCGCTCTGTACATAAAGAGGAAATCACTGCGAGC 357
 Qy 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
 Db 358 GAGCCTCTGATCGAAGGCTCATGTCGCCATGTCCTCATTTGGGGCAACAGTCATCACA 417
 Qy 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
 Db 418 ACATCTCGGGGGCAAGGGCTGCTGCTGTGTAGAGGGCCCATGCTGCTCGCTCGCT 477

136 IleuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu 155
478 GTCTCTTACCTGTCAGTGGGCTGGTATGCTTTGGCGCCAATTGTGTACACTTTGCTC 537
156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
538 CTCGCAAGGCTCATTGATGGTTTCGGTATCGGTTTGGCGGTACACTTGTCTCTCTAC 597
176 IleSerGluIleAlaProSerGluIleArg--GlyLeuLeuAsnThrLeuProGlnPhe 194
598 ATCTCCGAAATCGACCGCACAGANATCTTGGGGCTGNTNGAACACAGTTTCCGCGAGTTC 657
195 SerGly--SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213
658 ATTGGGCTCAGGAGGATGTTCTCTCTACTGCAATGGTGTGGATGTCCTCATG 717
214 ProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233
718 CCCAAACCTGATGGAGGCTCATGCTGGAGTTCTGTGATCCCGTCACCTTATNACTTT 777
234 GlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerGlyArgMet 253
778 GGACTGACTGTCTTCTACTTTCCTGTAATCACCAGGTGGCTTGTNAGCAAGGAGATG 837
254 AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeu 273
838 GCGAGCGCAAGAGAGTGTGCAAGGCTCGCGGGAAGAGAGATGCTCANGGGAGANG 897
274 SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluTyrIleIle 293
898 GCTCTTCTAGTTGAAGGTTTGGGGGTCGTAAGATACACGATATTTNAGAGTACATCAT 957
294 GlyProAlaThrGluAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThr 313
958 GGACCTGCCACCGAGCGCGATGATCTGTGAATGACGCGTATTAAGNACAAATCACA 1017
314 LeuTyrGlyProGluGluGlnSerTrpIleAlaArgProSerLysGlyProIleMet 333
1018 CTTTATGGGCTGAAGAAGCCAGTCATGATGCTCGAATTCCTTAAGGGACCCATCATG 1077
334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
1078 CTTGGAAGTGTCTTCTCTGATCTCGTCATGGGAGCATGTGTAACGAGTGTACCC 1137
354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
1138 CTTATGGATCCGATTGTGACACTTTTGTGTAGTGTCCATGAGAAATATGCTCAAGCTGGA 1197
374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
1198 GSAAGTATAGGAGCACATTTGTTCCAAACTTTGGAAAGTATGTTCAGTGTACAGATCAG 1257
394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413
1258 CATGCCAAAATGAGCAGTGGGATGAAGAAATCTTCATAGGATGACGAGGAGTACGA 1317
414 SerAspGlyAlaGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
1318 TCTGATGGTCAGAGGTGCTATGAGGACAACTCTCCATAGCCCATTTGTCTCAGGCGAG 1377
434 AlaThrGluValaGluGlyLysAspIleValHisGlyHisArgGlySerAlaLeuSer 453
1378 GCACAGGTCGGAAGGAGGACATTTGTGCACATGGTCACCTGGAGTGTCTTTGAGC 1437
454 MetArgArgGlnThrLeuLeuGlyGlyGlyAspGlyValSerSerThrAspIleGly 473
1438 ATGAGMAGGCAAGCCCTCTTAGGGAGGGTGGAGATGGTGTGAGCAGCACTGTATCGGT 1497
474 GlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGlu 493
1498 GGGGATGGCAGCTTGTCTGGAATGTCAGAGAGGAGGAGTGTGAGAAATGTTAGAAAGGAA 1557

494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer 513
1558 GGTGTTTCAAAGAGTCTACTTGCACCAAGAGGAGTCTCTGGCTCAAGAGGGCTCA 1617
514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533
1618 ATTGTTTCACTCCCGTGGTGGCGATGTTCTGAGGGTAGTAGTGTGTGATCATGCTGCT 1677
534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553
1678 GCTTTAGTAAGTCAGTCAGCACTTTTCTCAAAGGGTCTTGTGAACCAACGATGTCAGAT 1737
554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe 573
1738 GCTGCCATGGTTTCAACCATCTGAGGTAGCTGCCAAGGTTCCACGTTCGAAAGATTGTTT 1797
574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleGlnGlnPhe 593
1798 GAACCTGGAGTAGGCGTCCCTGTGTAGTCGGTGTGGAATTCAGATCCTTTCAACAGTTT 1857
594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613
1858 GCTGGAATTAACGGTGTCTGTACTATACCCACAAATCTTGTAGCAAGCTGGTGGCA 1917
614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
1918 GTTATTTCTTCCAAATTTGGTCTCAGCTCGGCATCAGCATCATCTTGTATCAGTTCTCTC 1977
634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
1978 ACTACTTACTAATGCTTCCCTTGCATTTGGCTTTCGATGCTGCTTATGGATCTTTCCGGA 2037
654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
2038 AGAAGGTTTTGCTGTAGGCACAAATCCAAATCTTGTATGATCTCTAGTTATCCTGTT 2097
674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
2098 GTGTCCAATCTAAATGTATTTGGGTACACTAGCCATGCTTTGCTCTCCACCATCAGTGT 2157
694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
2158 ATCGTCTACTTCTGCTGCTTGGTATGGGATTTGGTCCCATCCCAACATTTATGTGCA 2217
714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
2218 GAGATCTTTCCAAACAGGCTTCGTGGCTCTGTATGGCAATTTGTGCTTTTACATCTCG 2277
734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
2278 ATCGGAGATATCATCGTCACCTACAGCCTTCTGTGATGCTGAATGCTATTGGACTGGCG 2337
754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773
2338 GGTGTTTTTCAGCATATATGCGATGCTTGTGATTTCTTGTGTTGCTTCTCTCTCTTAAG 2397
774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPheAlaValGlyAla 793
2398 GTCCCTGAGACAAAGGGATGCCCTTGAAGTTATACCGAATCTTTTGCAGTGGTGGC 2457
794 LysGlnAlaAlaAlaLysAla 800
2458 AAGCAGCGGCTGCAAAAGCC 2478
RESULT 5
ADG47905
ID ADG47905 standard; cDNA; 2824 BP.
XX
AC ADG47905;
XX
DT 11-MAR-2004 (first entry)
XX
DE Corn Arabidopsis-like sugar transport protein cDNA #1.

XX Arabidopsis-like sugar transport protein;
 KW Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;
 KW corn; plant; gene; ss.
 XX Zea mays.
 OS
 XX Key Location/Qualifiers
 PH 238..2481
 FT CDS
 FT /product= "Corn Arabidopsis-like sugar transport protein"
 FT /transl_except= (pos:622..624, aa:Xaa)
 FT /note= "Xaa corresponds to an unknown amino acid"
 FT /transl_except= (pos:634..636, aa:Xaa)
 FT /note= "Xaa corresponds to an unknown amino acid"
 FT /transl_except= (pos:637..639, aa:Xaa)
 FT /note= "Xaa corresponds to an unknown amino acid"
 FT /transl_except= (pos:667..669, aa:Xaa)
 FT /note= "Xaa corresponds to an unknown amino acid"
 FT /transl_except= (pos:769..771, aa:Xaa)
 FT /note= "Xaa corresponds to an unknown amino acid"
 FT /transl_except= (pos:820..822, aa:Val)
 FT /transl_except= (pos:856..858, aa:Xaa)
 FT /note= "Xaa corresponds to an unknown amino acid"
 FT /transl_except= (pos:889..891, aa:Xaa)
 FT /note= "Xaa corresponds to an unknown amino acid"
 FT /transl_except= (pos:895..897, aa:Xaa)
 FT /note= "Xaa corresponds to an unknown amino acid"
 FT /transl_except= (pos:943..945, aa:Xaa)
 FT /note= "Xaa corresponds to an unknown amino acid"
 XX US2002199217-A1.
 XX 26-DEC-2002.
 XX 17-JAN-2002; 2002US-00051909.
 XX WPI; 2004-040967/04.
 XX P-PSDB; ADG47906.
 XX New isolated polynucleotide encoding a polypeptide having sugar transport
 XX protein activity, for producing a transformed plant and for use as probes
 XX in physical mapping.
 XX Disclosure; SEQ ID NO 1; 71bp; English.
 XX The invention relates to Arabidopsis-like or Beta-vulgaris-like sugar
 XX transport proteins and their corresponding nucleic acid sequences. The
 XX sequences of the invention are useful to transform a cell. These are also
 XX useful to produce a transgenic plant. Probes derived from sequences
 XX encoding sugar transport protein may be used for physical mapping. The
 XX present sequence is corn Arabidopsis-like sugar transport protein cDNA.
 XX Sequence 2824 BP; 644 A; 649 C; 745 G; 775 T; 0 U; 11 Other;
 SQ
 Alignment Scores:
 Pred. No.: 4,19e-298 Length: 2824
 Score: 3517.00 Matches: 684
 Percent Similarity: 95.18% Conservative: 27
 Best Local Similarity: 91.57% Mismatches: 34
 Query Match: 86.31% Indels: 2
 DB: 12 Gaps: 2
 US-10-051-909-32 (1-800) x ADG47905 (1-2824)
 QY 56 MetSerGlyAlaValLeuAlaValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75

Db 238 ATGGGGGGCCGCTGATGGTCGCATCGCGCTCTATCGCACTTGTCTGCGAGGCTGG 297
 QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleIleLysGluPheGlnLeuGlnAsn 95
 Db 298 GACAAATGCGACAATTGCTGGAGCGCTCTCTACATAAAGAAGAAATTCACCTGCGAGC 357
 QY 96 GluProThrValGluGlyLeuValSerMetSerLeuIleGlyAlaThrIleValThr 115
 Db 358 GAGCCTCTGATCGAAGGCTCATCGTCGCAATGTTCTTATTTGGGCAACAGTCATCACA 417
 QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgProMetLeuIleLeuSerSer 135
 Db 418 ACATCTCCGGGGCCAGGGCTGACTGGTGGTAGGAGGCCATGCTGGTCTGGCTGGCT 477
 QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrProAsnValTyrValLeuLeu 155
 Db 478 GTCTCTACTTCGTCAGTGGGCTGGTATGCTTTGGGCGCAATTGGTACATCTTGCTC 537
 QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
 Db 538 CTGCAAGGCTCATGATGGTTCGGTATCGGTTGGCGGTACACATGTTCTCTCTAC 597
 QY 176 IleSerGluIleAlaProSerGluIleArg---GlyLeuLeuAsnThrLeuProGlnPhe 194
 Db 598 ATCTCCGAAACTGCAACGCGACAGANATTCTTGGGGCTGTNGAACACAGTTGCCGAGTTC 657
 QY 195 SerGly---SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213
 Db 658 ATTGGGTGAGGAGGAGTGTCTCTCTACTGTCATGTTGGGATGTCCTCATG 717
 QY 214 ProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233
 Db 718 CCCAAACCTGATGGAGGCTCATGCTTGGAGTTCGTGCGATCCCGTCACTTAINTACTTT 777
 QY 234 GlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerGlyArgMet 253
 Db 778 GGACTGACTGTCTTCTACTTCTGCTGAATCACCAAGTGGCTTGTNAGCAAGAAGAGATG 837
 QY 254 AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeu 273
 Db 838 GCGGAGGCGAGAGAGTGTGCAAGGCTCGCGGAAGAGAGATGTCTCANGGAGANG 897
 QY 274 SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleLe 293
 Db 898 GCTCTCTAGTTGAAGGTTTGGGGTCTCGTAAGATACACGTATTTNAGAGTACATCAT 957
 QY 294 GlyProAlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGlnIleThr 313
 Db 958 GGACTGCCACCGAGGAGCGCGATGATCTTGTAACTGATAGGAACAAATCACA 1017
 QY 314 LeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMet 333
 Db 1018 CTTTATGGGCTGAAGAAGCGCAGTCATGATGATGCTCGACCTCTTAAGGAGCCCATCATG 1077
 QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
 Db 1078 CTTGGAAGTGTGCTTCTCTGTCATCTCGTCATGGGAGCATGTGTAACAGAGTACACC 1137
 QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
 Db 1138 CTTATGGATCCGATGTGACACTTTTGGTAGTGTCCATGAGAATATGCTCAAGTGA 1197
 QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
 Db 1198 GGAAGTATGAGGACACATTTGTTCCAAACTTTTGGAGTATGTTCAGTGTCAACAGATCAG 1257
 QY 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413
 Db 1258 CATGCCAAATATGAGCAGTGGATGAAGAGAAATCTTCATAGGATGACAGGAGTACGCA 1317
 QY 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433

Db	1318	TCTGATGGTGCAGGAGTCACTATGAGGACAATCTCCATAGCCCATGCTGTCCAGGCGAG	1377
Qy	434	AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer	453
Db	1378	GCAACAGGTGCGAAGGAGGACATTTGTGCACCATGTTCCCGTGGAGTCTTTGAGC	1437
Qy	454	MetArgArgGlnThrLeuLeuGlyGlyGlyAspGlyValSerSerThrAspIleGly	473
Db	1438	ATGAGAAGCAAGACCTCTTAGGGGAGGTGGAGATGGTGTGAGCAGCACTGATATCGGT	1497
Qy	474	GlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGlu	493
Db	1498	GGGGGATGACGCTTCTTGGAAATGGTCAGAGAGGAGGTGAGATGGTGAAGAAGNA	1557
Qy	494	GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer	513
Db	1558	GGTGGTTTCAAAAGAGTCTACTTGCCACCAAGAGGGAGTTCCTGGCTCAAGAGAGGGCTCA	1617
Qy	514	IleValSerLeuProGlyGlyAspValPheGluGlySerGluPheValHisAlaAla	533
Db	1618	ATTGTTTCTACTTCCCGGTGGCGATGTTCTTGGGGTAGTGAGTTTGATCATGCTGCT	1677
Qy	534	AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp	553
Db	1678	GCCTTTAGTAGTCAGTCAGCACATTTCTCAAGGGTCTTGTCTGAACCGCATGTCAGAT	1737
Qy	554	AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe	573
Db	1738	GCTGCCATGGTTCCACCATCTGAGGTAGTGCACAAAGGTTTCAGTTGGAAAGATTGTGTT	1797
Qy	574	GluProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPhe	593
Db	1798	GAACCTGGAGTGAGGGTGCCTGTTAGTCGGTGTGGATTTCAGATCCTTCAACAGATT	1857
Qy	594	AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluAlaGlyValAla	613
Db	1858	GCTGGAATAAACGGTGTCTGTACTATACCCACAAATTTCTGAGCAAGCTGGTGGCA	1917
Qy	614	ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerLeu	633
Db	1918	GTTATTCTTTCCAAATTTGCTCTCAGCTCGGATCAGATCCATCTTGATCAGTCTCTC	1977
Qy	634	ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly	653
Db	1978	ACTACCTTACTAATGCTTCTCTGCAATGGCTTTGCCATGCTGCTTATGGATCTTCCGGA	2037
Qy	654	ArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal	673
Db	2038	AGAAGGTTTTTGTGTAGCACAAATTCCAATCTTGATGCACTCTAGTCTATCCTGGTT	2097
Qy	674	ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal	693
Db	2098	GTGTCCAATCTAATGATTTGGGTACACTAGCCCATGCTTGTCTCCACCATCAGTGT	2157
Qy	694	IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla	713
Db	2158	ATCGTCTACTTCTGCTGCTTCTGTTATGGGATTTGGTCCCATCCCAACATTTTATGTGCA	2217
Qy	714	GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTyr	733
Db	2218	GAGATCTTTCCAAACAGGGTTCGTGGCTCTGATTTGCCATTTGTGCCCTTTACATCTCG	2277
Qy	734	IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla	753
Db	2278	ATCGGAGATATCATCGTCACACAGCCTTCCTGTGATGCTGAATGCTATTGGACTGGCG	2337
Qy	754	GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys	773
Db	2338	GGTGTGTTTCAGCATATATGAGTCGATGCTTGTATTTCTTTGTTGCTGCTTCCTTAAG	2397
Qy	774	ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla	793
Db	2398	GTCCCTGAGACAAAGGGATGCCCTTGGAGTTATTACCGAATTTCTTGCAGTTGGTGGC	2457

QY	794	LysGlnAlaAlaAlaLysAla	800
Db	2458	AAGCAAGCGGTGCAAAAGCC	2478
RESULT 6			
ABKS1965			
ID	ABKS1965	standard; cDNA; 2601 BP.	
XX	AC	ABKS1965;	
XX	27-AUG-2002	(first entry)	
DE	Soybean contig encoding A. thaliana-like sugar transport protein.		
XX	Soybean; Arabidopsis thaliana-like sugar transport protein;		
KW	carbohydrate transport; grain filling; annual field crop; plant; gene;		
KW	ss.		
OS	Glycine max.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	175..2388	
FT	/*tag= a		
FT	/product= "Soybean Arabidopsis thaliana-like sugar		
FT	transport protein"		
XX	US6383776-B1.		
PN			
XX			
PD	07-MAY-2002.		
XX			
PF	14-APR-1999; 99US-00291922.		
XX			
PR	24-APR-1998; 98US-0083044P.		
XX			
PA	(DUPO) DU PONT DE NEMOURS & CO E I.		
PI	Allen SM, Hitz WD, Kinney AJ, Tingey SV;		
XX			
DR	WPI; 2002-453364/48.		
DR	P-PSDB; AAU97204.		
XX			
PT	New nucleic acid encoding plant sugar-transport proteins, useful for		
PT	preparing transgenic plants with altered carbohydrate distribution.		
XX			
PS	Claim 3; Col 35-38; 54pp; English.		
XX			
CC	The present invention relates to the isolation of plant polynucleotide		
CC	sequences encoding an Arabidopsis thaliana-like sugar transport protein		
CC	or Beta vulgaris-like sugar transport protein. The polynucleotide		
CC	sequences are useful for altering the level of sugar transport protein		
CC	in plants, i.e. for control of carbohydrate transport and distribution		
CC	in plant cells, e.g. during grain filling of annual field crops (e.g. corn,		
CC	rice, soybeans, and wheat), and, for studying carbohydrate flows and		
CC	sugar transport. The polynucleotide sequences can also be used to isol		
CC	cDNA sequences and genes that encode homologues of the new proteins. T		
CC	present sequence representing a contig assembled from various soybean		
CC	cDNA clones encodes an Arabidopsis thaliana-like sugar transport prote		
XX			
SQ	Sequence 2601 BP; 648 A; 543 C; 644 G; 766 T; 0 U; 0 Other;		
Alignment Scores:			
Pred. No.:	3.54e-224	Length:	2601
Score:	2674.00	Matches:	522
Percent Similarity:	82.20%	Conservative:	92
Best Local Similarity:	69.88%	Mismatches:	115
Query Match:	65.62%	Indels:	18
DB:	6	Gaps:	8
US-10-051-909-32 (1-800) x ABKS1965 (1-2601)			
Qy	56	MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlyTyr	75

175 ATGAAGGTCCTGCTTGTGTTATTTGCGCTTCCATTTGTTAAATTTCTCCAAAGGATGG 234
Qy
76 AspAsnAlaThrIleAlaAlaValLeuTyrIleValGluPheGlnLeuGlnAsn 95
Db
235 GATAATGCTACCATCGCGGGCTAATGGTTACATTAAGAAGACCTCTCTTTGGGAACA 294
Qy
96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db
295 -----ACTATGAAAGGCTTGTGGTGGCATGCTCCTGATTGAGCAACGGTAAATCACC 348
Qy
116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
Db
349 ACATGCTCTGCTCTATAGCGGATTGCGTGGCGGACCCATGATGATATCTCATCT 408
Qy
136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrPheProAsnValTyrValLeuLeu 155
Db
409 GTGCTCTATTTCTTGGGTGGTTGGTGATGCTGTGTGCTCCCAATGTGTATGTGTGTC 468
Qy
156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
Db
469 TTGGCGAGGCTACTTGATGGAATTTGGGATTTGGCTTGTGACTCTTCTCCCGGTCTAT 528
Qy
176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
Db
529 ATATCTGAAACGCGCGCTCTGAATATAGGGGTGCTGTGAATACGCTTCTCAGTTTCA 588
Qy
196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
Db
589 GGCTCTGGAGCAATGTTTGTCTACTGTATGGTTTGGTATGTCATTTGAGTCCCGCG 648
Qy
216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
Db
649 CCTAGCTGGAGGCTCAGCTTGGGGTCTGTCTATTTCTCTCTTGTATTTTGCATTG 708
Qy
236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255
Db
709 ACCATTTTCTGCGCGAGTCTCTCGTGGTGGTCTGAGCAAGGAGGATGCTCGAG 768
Qy
256 AlaIysIysValLeuGlnIysLeuArgGlyLysAspValSerSerGlyLeuSerLeu 275
Db
769 GCTAAGAAGTGCTCCAAAGATTCCCGAAGAGGAGGATGTGTCAAGCGAGATGCGATTG 828
Qy
276 LeuLeuGluGlyLeuGluValGlyArgPheThrSerIleGluGluTyrIleIleGlyPro 295
Db
829 CTGGTTGAAGTCTCTGGGATTTGGGTGATACATCTATCGAAGATGATACATAATTTGG 888
Qy
296 AlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGluIleThrLeuTyr 315
Db
889 GCTGACGATGTGCTGATGTCATGAACATGCAACAGAGAAAGATAAATTCGATTATAT 948
Qy
316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335
Db
949 GGATCCCAAGCAGCCTTCTTGGTTATCAAAACCTGTCTCAGGACAGATTCATTGGC 1008
Qy
336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355
Db
1009 -----CTTGGTCCACACCATGGAAGCATCATCAACCAAGGATGCGCCCTCATG 1056
Qy
356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly----- 373
Db
1057 GATCCTCTGTGTGACATGTTGTGATGATTCATGAGAGCTCCCGAGACAGAGCAAGA 1116
Qy
374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db
1117 GGAAGCATCGAAGCACTCTGTTTCCAAATTTTGAAGCATGTTTCAGCACCTCCTCAGCG 1176
Qy
394 HisAlaIysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413
Db
1177 CATGCTAAATTTGAACAATGGGATGAAGAAAGCTTACAAGGGAACGTCGAGGACATC 1236
Qy
414 SerAspGlyAlaGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
Db
1237 TCAGATGCAACCGTGGGACTCCGATGATAATTTGCACAGTCTTTAATCTCACGCCAA 1296

Qy
434 AlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGlySerAlaLeu--- 452
Db
1297 ACAACAAGCCTTGA---AAAGACTTACCTCTCTCTCCATGGCAGTATCCTTGGC 1353
Qy
453 SerMetArgArgGlnThrLeuLeuGlyGluGly---GlyAspGlyValSerSerThrAsp 471
Db
1354 AGCATGAGGCTCAGACTAGTCTCATCAAGGGTCAAGTGAAGAGTGTGTATACAGGT 1413
Qy
472 IleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArg 491
Db
1414 ATTGGTGGTGGCTGGCACTGGCATGGAATGAGCTGATAAA---GGTAGAGTGGNAAA 1470
Qy
492 LysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArg 511
Db
1471 CAACAAGGAGGCTTAAAGAGATTTATTTACATGAGGAGGAGTTTCTGCATCTCGTCGT 1530
Qy
512 GlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHis 531
Db
1531 GGATCCATTGATCGATTCCCGTGAAGGC-----GAATTTGTCAG 1572
Qy
532 AlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMet 551
Db
1573 GCTGCTGCTTGGTAAGCCCAACCGCTCTTTACTCCAAAGGAGCTTATTGATGACACCCA 1632
Qy
552 SerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAsp 571
Db
1633 GTTGGGCTCGAATGGTTACCCATCTGACACAGCTTCAAAGGGGCCAAGTTGGAAGCT 1692
Qy
572 LeuPheGluProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln 591
Db
1693 CTCTCTGAAACGAGGGTTAGCATGCTGTTGTTGGAGTTGGAATACAAATACCTTTCAG 1752
Qy
592 GlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluAlaGly 611
Db
1753 CAGTTTTCAGGGATAAATGGGTTCTATATTACACACTCAAAATCTTGAAGGCGCGT 1812
Qy
612 ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer 631
Db
1813 GTTGAAGTTCTTCTTCAGATATAGGATTTGGCTCAGATCGGCATCATCTCTTATCAGT 1872
Qy
632 SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu 651
Db
1873 GCTTTCACACCTCTCTTGTGCTTCCCTGTATAGGGGTAGCCATGAAGCTCATGGATGTT 1932
Qy
652 SerGlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671
Db
1933 TCAGGCAAGAGCGAGTTGCTACTTACTACAAATCCCGTCTGATTGTGTGCATCATATT 1992
Qy
672 LeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691
Db
1993 TTGTCTATTGGAAGCCTGGTAAATTTTGGCAATTTGCCCATGCAGCAATCTCAACAGTA 2052
Qy
692 SerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu 711
Db
2053 TGGCTGTGTGTTATTTCTCTCTCTTGTGATGGGTATGAGCAATTCACAAATCCCTT 2112
Qy
712 CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731
Db
2113 TGCTCAGAGATTTTCCCACTAGGTTGGTGGCTCTCTGCATCTCTGTGCTATTAGTG 2172
Qy
732 PheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGly 751
Db
2173 TTCTGGATTGGAGACATCATCATCACTCTGCTGCTGTGATGCTCGCTCTTTAGGA 2232
Qy
752 LeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPhe 771
Db
2233 CTGTGGTGTGTATTCGCCATTTAGCAGTTGTTGTTTTCATCTCTGTTGGATTTTGTGTT 2292
Qy
772 LeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPheAlaVal 791
Db
2293 TTGAAGGTTCCAGAAACAAAGGCGCATGCCCTTTGAAGTCACTCTCTGAATTTCTTCTGTT 2352

QY	76	AspAsnAlaThrIleAlaAlaValLeuTyrIleLysGluPheGlnLeuGlnAsn	95
DB	235	GATAATGCTACCATCGCGGGCTAAATGGTTACATTAAAGAAAGACTTGCTTTGGGAACA	294
QY	96	GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr	115
DB	295	-----ACTATGGAAAGCCTTGGTGGCATGTCCCTGATTGGAGCAACGGTAATCACC	348
QY	116	ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer	135
DB	349	ACATGCTCTGGCTCTATAGCGGATTTGGCTCGGTCGGCGACCCATGATGATATCATCTCT	408
QY	136	IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu	155
DB	409	GTGCTCTATTCTTGGGCTGTTGGTGTGCTGTGGTCCCAAAAGTGTATGTGTGTGTC	468
QY	156	LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr	175
DB	469	TTGGCGAGGCTACTGTGATTTGGATTGGCATTTGGCTGTGACTCTTGTCCCGGTCTAT	528
QY	176	IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer	195
DB	529	ATATCTGAAACGGCGCGCTCTGAAATAAGGGGGTGGTGAATACGCTTCCTCAGTTCA	588
QY	196	GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer	215
DB	589	GGCTCTGGAGGAATGTTTTTGTCTACTGTATGGTTTTTGGCATGCTCATTCAGTCCCG	648
QY	216	ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu	235
DB	649	CCTAGCTGGAGGCTCATGCTTGGGGTCTGTCTATTCTCTCTCTCTATTTCGATTG	708
QY	236	ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu	255
DB	709	ACCATTTTTTCTTGGCCGAGTCTCTCGTGGCTGGTCAGCAAGGAAGATGCTCGAG	768
QY	256	AlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeu	275
DB	769	GCTAAGAAGTGCTCCAAACATTGCGCGAAGGAGGATGTGTACGGCGAGATGGCATTG	828
QY	276	LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro	295
DB	829	CTGGTTGAAGTCTCGGGATTCGGGGTGTATCATCTATCGAAGAGTACATTAATTTGGCC	888
QY	296	AlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr	315
DB	889	GCTGACGATGTGGCTGATGTGATGAACATGCAACAGAGAAGATAAAATTCGATTATAT	948
QY	316	GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly	335
DB	949	GGATCCCAACAGCGCTTCTTGGTATCAAAACCTGTCACTGTGACAGAGTTCTATTGGC	1008
QY	336	SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet	355
DB	1009	-----CTTGGCTCACACCATGGAACATCATCAACCAAAAGCATGCCCTCATG	1056
QY	356	AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly	373
DB	1057	GATCTCTGTGTACATGTTGTGACATTCATGAAGACTCCCGAGACAGAGCAAGA	1116
QY	374	GlySerMetArgSerThrLeuPheProAsnGlySerMetPheSerValThrAspGln	393
DB	1117	GGAAAGCATGCAAGCATCTGTTTCCAAATTTGGAAGCATGTTTCAGCACTGCTGAGCG	1176
QY	394	HisAlaLysAsnGluGlnTrpAspGluAsnLeuHisArgAspAspGluGluTyrAla	413
DB	1177	CATGCTAAATTTGAACAAATGGGATGAGAAAGCTTACAAAGGGAACTGAGACTACATG	1236
QY	414	SerAspGlyAlaGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln	433
DB	1237	TCAGATCAACCCGTGGGCACTCCGATGATAATTTGCACAGTCTCTTAAATCTCAACG	1296

QY 434 AlaThrGlyAlaGluGlyAspIleValHisGlyHisArgGlySerAlaLeu--- 452
 Db 1297 ACAACAGCCTTGAA---AAAGACTTACCTCCCTCCCTCCATGCCAGTATCCTTGGC 1353
 QY 453 SerMetArgGlnThrLeuLeuGlyGluGly---GlyAspGlyValSerSerThrAsp 471
 Db 1354 AGCATGAGGCGTCACTAGTCTCATGCAAGGTCAGGTGAGCAAGGTGGTAGTACAGGT 1413
 QY 472 IleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArg 491
 Db 1414 ATTGGTGTGGCTGCTGCACTGGCATGGAATGACTGATANA---GGTAGGATGAAA 1470
 QY 492 LysGluGlyGlyPheLysArgValTrpLeuHisGlnGluGlyValProGlySerArgArg 511
 Db 1471 CAACAAGAGGGGTTTAAAGGATTTATTTACATGAGGAGGAGTTTCTGTCATCTCGTCT 1530
 QY 512 GlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHis 531
 Db 1531 GGAATCATTTGATGCTTCCGCTGAGGC-----GAAATTTGCCAG 1572
 QY 532 AlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMet 551
 Db 1573 GCTGCTGCTTGGTAAGCAACCGCTCTTACTCCAAGGAGCTTATTGATGAGCACCCA 1632
 QY 552 SerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAsp 571
 Db 1633 GTTGGGCTGCAATGGTTTCAACCATCTGAGACAGCTTCAAGGGGCGCAAGTTGGAAGCT 1692
 QY 572 LeuPheGluProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln 591
 Db 1693 CTCTTGAACACCGGTTAAGCATGCTGTTGTTGGAGTTGGAATACAAATCTTCA 1752
 QY 592 GlnPheAlaGlyIleAsnGlyValLeuTyrrTrpProGlnIleLeuGluGlnAlaGly 611
 Db 1753 CAGTTTTCAGGGATAAATGGGTTCTATATTACACACCTTCAATCTTGAAGGCGCGT 1812
 QY 612 ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer 631
 Db 1813 GTTGAAGTCTCTTTTCAGATAGCATGGCTCAGACTCGCATCATCTCTTATCATG 1872
 QY 632 SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuMetAspLeu 651
 Db 1873 GCTTTCACACCTCTCTGATGCTTCCCTGATAGGCGTAGCCATGAAGCTCATGATGTT 1932
 QY 652 SerGlyArgArgPheLeuLeuGlyThrIleProIleLeuAlaSerLeuValIle 671
 Db 1933 TCAGGCAAGAGGAGTTGCTTACTCAATCCCGCTGCTGATGTGCTCATCTATTATT 1992
 QY 672 LeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691
 Db 1993 TTGGTCATTGGAAGCCTGTAATTTTGGCAATGTCGCCATGTCAGCAATCTCACAGTA 2052
 QY 692 SerValIleValTyrrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu 711
 Db 2053 TCGGTGTGGTTTATTTCTGCTGCTTGTGATGGGTTATGACCAATTTCCAAACATCTT 2112
 QY 712 CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731
 Db 2113 TGCTCAGAGATTTTCCCACTAGGTTGGTGGCTCTGCATTTGCTATCTGTGCTATTAGT 2172
 QY 732 PheTrpIleGlyAspIleIleValThrTyrrSerLeuProValMetLeuAsnAlaIleGly 751
 Db 2173 TTCTGATTGGAGATCATCATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2232
 QY 752 LeuAlaGlyValPheSerIleTyrrAlaValValCysLeuIleSerPheValPheValPhe 771
 Db 2233 CTGTGGTGGTATTTCGCATTTACGAGTTGTTTGTTCATCTCTGTGATATTGTGTTT 2292
 QY 772 LeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaVal 791
 Db 2293 TTGAAGGTTTCCAGAAACCAAGGCGATGCCCTTGAAGTCATCTGATTTCTTCTGTT 2352
 QY 792 GlyAlaLysGlnAlaAla 798

Db 2353 GGAGCAAGCAGGCTGCTTCT 2373
 RESULT 8
 ADG47911
 ID ADG47911 standard; cDNA; 2601 BP.
 XX
 AC ADG47911;
 DT 11-MAR-2004 (first entry)
 XX
 DE Corn Arabidopsis-like sugar transport protein cDNA #2.
 XX
 KW Arabidopsis-like sugar transport protein;
 KW Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;
 KW corn; plant; gene; ss.
 XX
 OS Zea mays.
 XX
 FH Key Location/Qualifiers
 FT CDS 40..2388
 FT /*tag= a
 FT /product= "Corn Arabidopsis-like sugar transport protein
 FT
 XX US2002199217-A1.
 XX
 PD 26-DEC-2002.
 XX
 PF 17-JAN-2002; 2002US-00051909.
 XX
 PR 24-APR-1998; 98US-0083044P.
 PR 14-APR-1999; 99US-00291922.
 XX
 PA (HELE/) HELENTJARIIS T G.
 XX
 PI Helentjaris TG;
 XX
 DR WPI; 2004-040967/04.
 DR P-PSDB; ADG47912.
 XX
 PT New isolated polynucleotide encoding a polypeptide having sugar transport
 PT protein activity, for producing a transformed plant and for use as probes
 PT in physical mapping.
 XX
 PS Disclosure; SEQ ID NO 7; 71pp; English.
 XX
 CC The invention relates to Arabidopsis-like or Beta-vulgaris-like sugar
 CC transport proteins and their corresponding nucleic acid sequences. The
 CC sequences of the invention are useful to transform a cell. These are also
 CC useful to produce a transgenic plant. Probes derived from sequences
 CC encoding sugar transport protein may be used for physical mapping. The
 CC present sequence is corn Arabidopsis-like sugar transport protein cDNA.
 XX
 SQ Sequence 2601 BP; 648 A; 543 C; 644 G; 766 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 3 54e-224 Length: 2601
 Score: 2674.00 Matches: 522
 Percent Similarity: 82.20% Conservative: 92
 Best Local Similarity: 69.88% Mismatches: 115
 Query Match: 65.62% Indels: 18
 DB: 12 Gaps: 8
 US-10-051-909-32 (1-800) x ADG47911 (1-2601)
 QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
 Db 175 ATGAAAGGTGCGCTCTCTTGTCTATTCGCTTCCATTTGTAATTTCTCCCAAGGATGG 234
 QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrrIleLysGluPheGlnLeuGlnAsn 95
 Db 235 GATAATGCTACCATCGCCGGGCTAATGTTACATTAGAAAGACCTTCTTGGGAACA 294

Qy 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db 295 -----ACTATGGAAAGGCTTGGTGGGCAATGCTCTGATTGGAGCAAGCTAATCACC 348
Qy 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
Db 349 ACATGCTCTGGTCTCTATAGCGGATTGGCTGGCGGACCCCATGATGATATCTCATCT 408
Qy 136 IleLeuTyrrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrrValLeuLeu 155
Db 409 GTGCTCTATTTCTTGGGTGGTTTGGTGATGCTGTGGTCCCAAAATGTGTATGTGTGTGC 468
Qy 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrr 175
Db 469 TTGGCGAGGCTACTTGTATGGGATTGGGATTGGCTGTGACTCTTCTCCCGGTCTAT 528
Qy 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
Db 529 ATATCTGAAACGGCGCGCTCTGAAATAAGGGGTCTGTGAATACGCTTCTCTAGTTTCAGT 588
Qy 196 GlySerGlyGlyMetPheLeuSerTyrrCysMetValPheGlyMetSerLeuSerProSer 215
Db 589 GGCTCTGGAGGAATTTTTGTCTACTGTATGCTGTTTGGCATGTCATGAGTCCCGCG 648
Qy 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
Db 649 CCTAGCTGGAGGCTCATGCTTGGGTCTCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCT 708
Qy 236 ThrIlePheTyrrLeuProGluSerProArgTrpLeuValSerIysGlyArgMetAlaGlu 255
Db 709 ACCATTTTTTCTTGTCCCGAGTCTCTCGTGGCTGGTCAGCAAGAGGATGCTCGAG 768
Qy 256 AlaIysIysValLeuGlnIysLeuArgGlyIysAspAspValSerGlyGlnLeuSerLeu 275
Db 769 GCTAAGAAAGTGTCTCCAAAGATTGCGCGAAGGAGGATGTGTCAGCGAGAGTGGCAATG 828
Qy 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluIleGlyPro 295
Db 829 CTGCTTGAAGTCTCGGGATTGGGGTGATACATCTATCGAAGGTACATAATTTGCCCT 888
Qy 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspIysGluGlnIleThrLeuTyrr 315
Db 889 GCTGACGATGGGCTGATGGTCAACATGCAACAGAAAGATAAAATTCGATTATAT 948
Qy 316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerIysGlyProIleMetLeuGly 335
Db 949 GGATCCCAAGCAGCGCTTCTTGGTTATCAAAACCTCTCACTGCACAGAGTTCTATTTGC 1008
Qy 336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355
Db 1009 -----CTTGGCTCACACCATGGAAGCATCATCAACCAAGCATGCCCTCATG 1056
Qy 356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly----- 373
Db 1057 GATCCTCTGTGTACACTGTTGGTAGCATTCATGAGAAGCTCCCGAGACAGAGCAAGA 1116
Qy 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db 1117 GGAAGATGCGAAGCACTCTGTTTCCAAATTTTGGAGAGATGTCACACTGCTGAGCCG 1176
Qy 394 HisAlaIysAsnGluGlnTrpAspGluAsnLeuHisArgAspAspGluGluTyrrAla 413
Db 1177 CATGCTAAATTTGAACATGGATGGAAGAGCTTACAAAGGAAACGTGAGGACTACATG 1236
Qy 414 SerAspGlyAlaGlyGlyAspTyrrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
Db 1237 TCAGATGCAACCCGTGGGAGTCCGATGATTAATTTGCACAGTCTCTTAATCTCAGCCAA 1296
Qy 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeu--- 452
Db 1297 ACAACAGCCCTTGA---AAGACTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1353

Qy 453 SerMetArgArgGlnThrLeuLeuGlyGluGly---GlyAspGlyValSerSerThrAsp 471
Db 1354 AGCATGAGGGCTCACAGTAGTCTCATCAAGGGGTCAGGTGAGCAAGGTGGTAGTACAGGT 1413
Qy 472 IleGlyGlyGlyTrpGlnLeuAlaTrpIysTrpSerGluIysGluGlyGluAsnGlyArg 491
Db 1414 ATTGGTGGTGGCTGGCAACTGGCATGGAATGGAATGCAATAAA---GGTGGAGATGAAAA 1470
Qy 492 LysGluGlyGlyPheIysArgValTyrrLeuHisGlnGluGlyValProGlySerArgArg 511
Db 1471 CAACAAGGAGGGTTTAAAGAGATTATTTATCATGAGGAGGAGTTTCTGCACTCTCGTCGT 1530
Qy 512 GlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHis 531
Db 1531 GGTATCCATTGTATGATCCCGGTGAAGGC-----GAATTTGTCCAG 1572
Qy 532 AlaAlaAlaLeuValSerGlnSerAlaLeuPheSerIysGlyLeuAlaGluProArgMet 551
Db 1573 GCTGCTGCTTGGTAAGCAACCGCTCTTTTACTTCAAGGAGCTTATTGATGGACACCCA 1632
Qy 552 SerAspAlaAlaMetValHisProSerGluValAlaAlaIysGlySerArgTrpIysAsp 571
Db 1633 GTTGGGCTGCAATGGTTTCAACCATCTGAGACACACTTCAAGGGGCCAAGTTGGAAGCT 1692
Qy 572 LeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln 591
Db 1693 CTCTCTTGAACCAAGGGTTTAAAGCATGCTTGTGTTGAGTTGGAATACAAATCTTCCAG 1752
Qy 592 GlnPheAlaGlyIleAsnGlyValLeuTyrrTyrrProGlnIleLeuGluGlnAlaGly 611
Db 1753 CAGTTTTCAGGGATAAATGGGGTTCTATATTACACACTCAAACTCTTGAAGAGCCCGT 1812
Qy 612 ValAlaValIleLeuSerIysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer 631
Db 1813 GTTGAAGTTCTTCTTTCAGATATAGGCATTCGCTCAGAGTCGGCATCTCTCTATCAGT 1872
Qy 632 SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu 651
Db 1873 GCTTTCACAACTTCTTGTATGCTTCCCTGTATAGCGGTAGCATGAAGCTCATGATGTT 1932
Qy 652 SerGlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671
Db 1933 TCAGCGAAGGACGATGCTTACTTACTAATCCCGCTGCTGATGTGCTCACTCATATT 1992
Qy 672 LeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691
Db 1993 TTGGTCAITGGAGCGCTGGTAAATTTTGGCAATGTGCGCCATTCGACAACTCTCAACAGTA 2052
Qy 692 SerValIleValTyrrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu 711
Db 2053 TCGCTGTGTGTATTCTTCTGCTCTTGTGATGGGTTATGGACCAATTCCAAACATCCTT 2112
Qy 712 CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731
Db 2113 TGCTCAGAGATTTTCCCACTAGGGTGGTGGCTCTGCAATTCGCTATCTGTGCATTAGTG 2172
Qy 732 PheTrpIleGlyAspIleIleValThrTyrrSerLeuProValMetLeuAsnAlaIleGly 751
Db 2173 TTCTGGATTGGAGACATCATCATCATCTCGCTCGCTGTGATGCTCGGCTCTTCTTCTT 2232
Qy 752 LeuAlaGlyValPheSerIleTyrrAlaValValCysLeuIleSerPheValPheValPhe 771
Db 2233 CTGTGTGTGTGATTTGCCATTTACGAGTTGTTGTTGTTTCATCTCGTGGATATTTGTGTTT 2292
Qy 772 LeuIysValProGluThrIysGlyMetProLeuGluValIleThrGluPhePheAlaVal 791
Db 2293 TTGAAGGTTCCAGAAACAAAGGCGCATGCCCTTGAAGTCATCTCTGAAATCTTCTTCTGTT 2352
Qy 792 GlyAlaIysGlnAlaAlaAla 798
Db 2353 GGAGCAAGAGCGCTCTTCT 2373

155 LeuLeuAlaArgPheValAspGlyPheGlyLeuAlaValThrLeuValProLeu 174
 292 TGCTTCTAGGCTTCTTAATGGTTGGTGGCGGCTCGGGTTACATTTGCCCTGTT 351
 175 TyrSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPhe 194
 352 TACATTCTGAACCGCTCTCCGGAGATCAGAGGACAGTAAATACTCTCCCTCAGTTT 411
 195 SerGlySerGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerPro 214
 412 CTGGCTCTCGTGGAAATGTTTGTCTACTGATGTTTTCATGTCCTCGAGTGAC 471
 215 SerProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGly 234
 472 TCCCTAGCTGGAGAGCCATGCTCGGTGCTCTCGATCCCTCTCTTTATTGTTT 531
 235 LeuThrIlePheTyrLeuProGluSerProArgTyrLeuValSerIleGlyArgMetAla 254
 532 CTCACGGTGTATTATTTGCCCGAGTCTCTCGTGGCTGTGTAGTAAAGAGAAATGGAC 591
 255 GluAlaIleLysValLeuGlnLysLeuArgGlyLysAspValSerGlyLeuLeuSer 274
 592 GAGCTAAGCGAGTCTTCAACAGTTATGTGGAGAGAGATGTTACCGATGAGATGGCT 651
 275 LeuLeuLeuGlyLeuGluValGlyGlyAspThrSerIleGluIleIleGly 294
 652 TTACTAGTGAAGACTAGATATAGGAGGAGAGAAAAAATGGAAGATCTCTTAGTAAC 711
 295 ProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGlu---GlnIleThr 313
 712 TTGAGAGATCATGAAGTGTATGATACACTTGAACCGTGTGAGAGATGACAAATGGCG 771
 314 LeuTyrGlyProGluGluGlyGlnSerTyrIleAlaArgProSerLysGlyProIleMet 333
 772 CTATTGAACCCAGAGATCAATGCTACCTGTCTGATGACCTGTCACAGAACAA----- 825
 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
 826 ---AATAGCTCCTTGGGCTACGCTCGCCACGGAAGCTTAGCAACCAAGCATGATC 882
 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
 883 CTTAAGATCCGCTCGTCAATCTTTTGGCAGTCTCCACGAGAAGATGCCAGAGCAGGC 942
 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
 943 GGAAACACTCGGAGTGGGATTTCCCTCATTTCCGAAGCATGTTCTAGTACTACTGCCGAT 1002
 394 -----HisAlaLysAsnGluGlnTyrAsp-----GluGluAsnLeuHisArgAsp 408
 1003 CGCGCTCAGCGTAAACCGGCTCATTTGGAAAAGACATAGAGAGCCATTACACAAAGAC 1062
 409 AspGluGluTyrAlaSerAspGlyAlaGlyAspTyrGluAsp-----AsnLeu 425
 1063 AATGATGACTATGCGACTGATGATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1122
 426 HisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluLysAspIleValHisHis 445
 1123 CGTAGCCCCCTTAATGTCGCCCGACACCAAGCATGAGC---AAGGATATGATCCACAT 1179
 446 GlyHisArgGlySerAlaLeuSerMetArgGlnThrLeuLeuGlyGluGlyAsp 465
 1180 CCTACAAGTGGAGCACTTTAAGCATGAGACGACACAGTACGCTTATGCAA---GGCAAC 1236
 466 GlyValSerSerThrAspIleGlyGlyTyrGlnLeuAlaTyrIlePheSerGluLys 485
 1237 GCGAAGATGACATGGGAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1296
 486 GluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGly 505
 1297 GAA-----TACAGAGGATTTATCTTAAAGAGATGGA 1329

506 ValProGlySerArgArgGlySerIleValSerLeuProGlyGlyAspValPheGlu 525
 1330 GCTGAA---TCTGCCGTGGCTCGCATCTCTATTCGGAGGTCGGAT-----GGT 1380
 526 GlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGly 545
 1381 GGAGCGACTACATTCACGCTTCTGCCCTTGTAGCAGATCTGTTCTTGGTCTAAATCA 1440
 546 LeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLys 565
 1441 -----GTTTCATGGATCGCCATGTTCCCGGAGAGAAATGTCGCTCT 1485
 566 GlySerArgTyrPheLysAspLeuPheGluProGlyValArgAlaLeuLeuValGlyVal 585
 1486 GGACCACTCTGCTGCTCTCTTGAACCTGGTGTAAAGCGTGCCTTGGTGTGGTGTGTC 1545
 586 GlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGln 605
 1546 GGCATTCAATACATGACAGCTTTTTCAGGTATCAATGGAGTTCTCTACTACATCTCAG 1605
 606 IleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSer 625
 1606 ATTCTCGAAACGGGCTGGCGTAGATATTCTTTCGAGCCTCGGACTAAGTTCATCTCT 1665
 626 AlaSerIleLeuIleSerSerLeuThrLeuLeuMetLeuProCysIleGlyPheAla 645
 1666 GCGTCATTCTCATCAGCGGTTTAAACAACTTACTCATGCTCCCGCATTTGCTGTC 1725
 646 MetLeuLeuMetAspLeuSerGlyArgPheLeuLeuLeuGlyThrIleProIleLeu 665
 1726 ATGAGACTCATGATGATCGGAAGAGGTCAATTACTTCTCGACAAATCCACAGTTCTC 1785
 666 IleAlaSerLeuValIleLeuValValSerAsnLeuLeuLeuGlyThrLeuAlaHis 685
 1786 ATTGTCTCATTGTCGTCTGTGTCATCAGCGAGCTCATCCACATCAGCAAAAGTCGTAAC 1845
 686 AlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGly 705
 1846 CGACACTCTCCACAGGTTGTGCTGCTCTACTTCTTCTGTCGATGGGTACGGT 1905
 706 ProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIle 725
 1906 CCCATTCCAAACATCTCTGTCTGTAATCTTCCCAACAGAGTCGCTGCTCTGTCATC 1965
 726 AlaIleCysAlaPheThrPheTyrIleGlyAspIleIleValThrTyrSerLeuProVal 745
 1966 GCCATATGCTGATGCTGCTTTTGGATTGGAGACATTAATGTACGACTACTACTCCCGTT 2025
 746 MetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIle 765
 2026 CTCCTCAGCTCGATCGGACTAGTTGGTGTGTTTACGATTTACGCTGCGGTTGCGTTATC 2085
 766 SerPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIle 785
 2086 TCATGGATCTCTGTTTACATGAAGTCCCGGAGACTAAAGCATGCCCTTTGGAAGTTATC 2145
 786 ThrGluPhePheAlaValGlyValLysGlnAlaAlaLysAla 800
 2146 ACAGACTACTTTGCTTTGGAGCT---CAAGCTCAAGCTTCTGCT 2187

RESULT 11

ADG87978

ID ADG87978 standard; cDNA; 2205 BP.

XX AC ADG87978;

XX DT 22-APR-2004 (first entry)

DE A. thaliana RPP4-upregulated pathogen infection-related gene #420.

XX pathogen infection-related gene; plant; Peronospora parasitica;

XX defence mechanism; RPP4; pathogen resistance; transgenic plant; oomycete;

XX fungus; bacterium; virus; nematode; insect; aphid; gene; ss.

Db 81 GTTGCTGGACCAAAATCTGTGGC-----CTTGATCTAGGAAAGGAAGCATG 128
 Qy 348 ValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGlu 367
 Db 129 GCAAAATCCAAAGC--AGTCAGTGGACCCCTCTAGTACCCCTCTTTGGTAGGTACATGAG 185
 Qy 368 AsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPhedGlySerMet 387
 Db 186 AAGCTCCAGAAACAGGA-----AGCACCCCTTTTCCACACTTTGGGAGATG 233
 Qy 388 PheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArg 407
 Db 234 TTCAGTGTGGGGAAATCAGCAAGGAATGAAGATGGGATGAGAAAGCCCTAGCCAGA 293
 Qy 408 AspAspGluGlnTrpAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSer 427
 Db 294 GAGGGTGATGATTATGCTCTGAT-----GCTGGTGATTTCTGATGACAAATTTGCAGAGT 347
 Qy 428 ProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisGlyHis 447
 Db 348 CCATTGATCTCAGCTCAACCAACAGCTGGAT---AAGACATACCTCTCATGCCCAT 404
 Qy 448 ArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu---GlyGluGlyGlyAspGly 466
 Db 405 AGTAACCTTGCA---AGCATGAGGCAAGGTAGTCTTTTACATGAAATTCAGGAGAACCC 461
 Qy 467 ValSerSerThrAspIleGlyGlyGlyTrpGluLeuAlaTrpLysTrpSerGluLysGlu 486
 Db 462 ACTGGTAGTACTGGGATTTGGTGGTTGGCAGCTAGCATGGAATGGTCTGAAAGAGAG 521
 Qy 487 GlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTrpLeuHisGlnGluGlyVal 506
 Db 522 GGCCAGATGAAAGAGGAAGGTGGCTTCAAGAGAATATATTACCAAGATGTTGTTG 581
 Qy 507 ProGlySerArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGly 526
 Db 582 TCTGGATCTAGACGTGGTCTGGTTTCACTCCCT---GGCGGTGATTTACCAACTGAC 638
 Qy 527 SerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeu 546
 Db 639 AGTGAGTTGTACAGGCTGTCTCTGGTGAGTCAGCTGCCCTTTTATATGAGGACCTT 698
 Qy 547 AlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaLysGly 566
 Db 699 ATGCGTCAACGGCAGTTGGACAGCATATGATTCATCCCTCTGAAACAAATTCGAAAGGG 758
 Qy 567 SerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGly 586
 Db 759 CCAAGTTGGAGTATCTTTTGAACCTGGGGTGAAGCATGATGATTTGGGGGTGGGA 818
 Qy 587 IleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTrpTrpThrProGlnIle 606
 Db 819 ATGCAAAATCTTCAGCAGTTCTCTGGTATAAATGGGGTCTCTACTATACGCCCTCAAT 878
 Qy 607 LeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAla 626
 Db 879 CTTGAGCAGCAGGTGTGTATCTCTTTCAAGCCTAGGCTTGGTCTACTTCTTCA 938
 Qy 627 SerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet 646
 Db 939 TCCTTTCTTATTAGTCGGTGACACCTTCTTGATGCTTCTGTATAGCCATGGCCATG 998
 Qy 647 LeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIle 666
 Db 999 AGGCTCATGTGATATTTTCAGGCAAGAGGACCTTCTGCTCAGTACAAATCCCGCTCAATA 1058
 Qy 667 AlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAla 686
 Db 1059 GCAGCTCTTCTCATATTAGTCTCGGAGAGCTTGTGGATTTGGATCCACTGCAATGCA 1118
 Qy 687 LeuLeuSerThrValSerValIleValTrpPheCysCysPheValMetGlyPheGlyPro 706
 Db 1119 TCAATCTCAACCATTAGTGTATTGTCTATTCTTCTTCTTCTTCTGATGGGATTTGGACCA 1178

Qy 707 IleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAla 726
 Db 1179 ATTCCTATATACTTTGTGCAGAGATCTTCCCACTCGAGTTCGTGGTCTCTGCATTGCT 1238
 Qy 727 IleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTrpSerLeuProValMet 746
 Db 1239 ATTGTGCCCTTACTTTTGGATCTGTGATATCATTTGTCACCTACACACTCCAGTTATG 1298
 Qy 747 LeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTrpAlaValValCysLeuIleSer 766
 Db 1299 CTCAAATCTGTAGGCTCGCTGGTGTGTTTGGTATTTATGCTGCTGCTGCATTATGCA 1358
 Qy 767 PheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThr 786
 Db 1359 TGGGTGTTTGTCTTTTGAAGTTCCAGAAACCAAGGCACTGCCACTGGAAGTATCAT 1418
 Qy 787 GluPhePheAlaValGlyAlaLysGln 795
 Db 1419 GAGTCTTCTCTGTCGGAGCAAAACAG 1445

RESULT 13
 ABX93202
 ID ABX93202 standard; cDNA; 1692 BP.
 XX
 AC ABX93202;
 DT 29-MAY-2003 (first entry)
 XX
 DE cDNA encoding soybean sugar transport protein #2.
 XX
 KW Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
 KW plant sugar transport protein; carbohydrate transport; soybean;
 KW carbohydrate distribution; plant; gene; ss.
 XX
 OS Glycine max.
 XX
 PN US2002178468-A1.
 XX
 PD 28-NOV-2002.
 XX
 PF 17-JAN-2002; 2002US-00051902.
 PR 24-APR-1998; 98US-0083044P.
 PR 14-APR-1999; 99US-00291922.
 XX
 PA (ALLEN/) ALLEN S M.
 PA (HITZ/) HITZ W D.
 PA (KINN/) KINNEY A J.
 PA (TING/) TINGEY S V.
 XX
 PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
 XX
 DR WPI: 2003-340957/32.
 DR P-PSDB; ABU08330.
 XX
 PT Novel plant sugar transport proteins and nucleic acid encoding the
 PT protein useful for producing transgenic plants having altered levels of
 PT sugar transport protein.
 XX
 PS Claim 2; Page 22-23; 56pp; English.
 CC
 CC The present invention relates to the isolation of Arabidopsis thaliana-
 CC like or Beta vulgaris-like sugar transport proteins, and the
 CC polynucleotide sequences encoding them. The plant sugar transport
 CC proteins of the invention have been isolated from corn, rice, soybean,
 CC and wheat. The polypeptides of the invention may be used for altering the
 CC level of expression of a sugar transport protein in a host cell, by
 CC transforming a host cell with a chimeric construct encoding all, or a
 CC portion of the sugar transport protein, in sense or antisense
 CC orientation. Particularly, the polypeptides may provide a means to
 CC control carbohydrate transport and distribution in plants. ABX93198-
 CC ABX93205 represent cDNA sequences encoding Arabidopsis thaliana-like

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 13, 2004, 11:54:50 ; Search time 169.106 Seconds
(without alignments)
3362.578 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSLAVQTPTDLDLR.....PLEVITEFFAVGAKQAATA 800

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/uspto.spool.h/US10051909/runat 13102004 123337 19931/app_query.fasta_1.1678
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=arni -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human4.0.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051909 @CGN 1 1 116 @runat 13102004 123337 19931 -NCFU=6 -ICFU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*
1: /cgn2_6/prodata/1/ina/5A COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
5: /cgn2_6/prodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3517	86.3	2824	3	US-09-291-922-1
2	2674	55.6	2601	3	US-09-291-922-7
3	1670.5	41.0	1692	3	US-09-291-922-9
4	1424	34.9	1487	3	US-09-291-922-13
5	943	23.1	1009	3	US-09-291-922-15
6	562	13.8	870	3	US-09-291-922-5
7	496	12.2	1853	3	US-09-291-922-23
8	494.5	12.1	2017	3	US-09-291-922-21
9	491.5	12.1	1914	3	US-09-291-922-19
10	489	12.0	1872	3	US-09-291-922-27
11	463.5	11.4	2089	3	US-09-291-922-25
12	447	11.0	1431	4	US-09-489-039A-4762

ALIGNMENTS

RESULT 1

US-09-291-922-1
; Sequence 1, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; EARLIER FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 2824
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (29)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (622)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (636)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (638)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (569)

13	440.5	10.8	1545	4	US-09-489-039A-4731	Sequence 4731, Ap
14	427	10.5	1506	4	US-09-489-039A-4560	Sequence 4560, Ap
15	426	10.5	1566	4	US-09-489-039A-2378	Sequence 2378, Ap
16	418.5	10.3	1668	4	US-09-614-221A-420	Sequence 420, App
17	412	10.1	441529	3	US-09-103-840A-2	Sequence 1, Appli
18	368	9.0	4403765	3	US-09-103-840A-1	Sequence 2, Appli
19	355	8.7	2856	4	US-09-643-597-135	Sequence 135, App
20	355	8.7	2856	4	US-09-480-884A-135	Sequence 135, App
21	355	8.7	2856	4	US-09-542-615A-135	Sequence 135, App
22	355	8.7	2856	4	US-09-606-421B-135	Sequence 135, App
23	355	8.7	2856	4	US-09-221-107-135	Sequence 135, App
24	355	8.7	2856	4	US-09-456-396A-135	Sequence 135, App
25	355	8.7	2856	4	US-09-476-496A-135	Sequence 135, App
26	355	8.7	2856	4	US-09-630-340B-135	Sequence 135, App
27	354.5	8.7	3000	2	US-08-928-692-9	Sequence 9, Appli
28	354.5	8.7	3000	3	US-09-339-972-9	Sequence 9, Appli
29	352.5	8.7	1626	4	US-09-614-221A-521	Sequence 521, App
30	346	8.5	5228	4	US-09-919-039-216	Sequence 216, App
31	343	8.4	1776	4	US-09-679-686B-11	Sequence 11, Appli
32	338.5	8.3	1943	4	US-09-774-528-168	Sequence 168, App
33	334.5	8.2	5227	4	US-09-919-172-79	Sequence 79, Appli
34	332.5	8.2	2592	3	US-09-591-025-8	Sequence 8, Appli
35	338.5	8.1	2592	4	US-09-894-927B-8	Sequence 8, Appli
36	322	7.9	1659	4	US-09-248-796A-6610	Sequence 6610, Ap
37	313.5	7.7	1704	4	US-09-614-221A-534	Sequence 534, App
38	312	7.7	1675	4	US-09-679-686B-1	Sequence 1, Appli
39	312	7.7	1695	4	US-09-614-221A-443	Sequence 443, App
40	307.5	7.5	1704	4	US-09-614-221A-96	Sequence 96, Appli
41	302	7.4	2298	4	US-09-248-796A-5965	Sequence 5965, Ap
42	301.5	7.4	3915	4	US-09-023-655-1104	Sequence 1104, Ap
43	293	7.2	987	4	US-09-710-279-1941	Sequence 1941, Ap
44	293	7.2	3065	4	US-09-710-279-3697	Sequence 3697, Ap
45	291.5	7.2	1752	4	US-09-679-686B-17	Sequence 17, Appli

QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
 Db 1978 ACTACCTTACTTAATGCTTCTTGGATTGGCTTTGCCATGCTGCTTATGGATCTTTCCGGA 2037
 QY 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuLeuLeuValIleLeuVal 673
 Db 2038 AGAAGGTITTTGCTGCTAGGACAAATCCAAATCTTGATAGCATCTCTAGTATCTCTGGTT 2097
 QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
 Db 2098 GFGTCCATCTAATTAATGTTGGTACACTAGSCCATGCTTGTCTCTCCACCATCAGTGT 2157
 QY 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
 Db 2158 ATCGCTACTTCTGCTGCTGCTATGGATTTGGTCCCATCCCAACATTTTATGACA 2217
 QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
 Db 2218 GAGATCTTTCCAAACAGGTTTCGTGGCTCTGTATGGCATTTGTGCTTTACATCTGG 2277
 QY 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
 Db 2278 ATCGAGATATCATGTCACCTACAGCCTTCTCTGTATGCTGAATGCTATTTGGACTGGG 2337
 QY 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773
 Db 2338 GGTGTTTTCAGATATATGCAATGCTAGTCTGTATGCTTTCCTTTGCTTCCTTAAG 2397
 QY 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
 Db 2398 GFCCTGTAGACAAAGGGATGCCCTTGAGTTATTACCGAAATCTTTGCAAGTGGTGG 2457
 QY 794 LysGlnAlaAlaAlaLysAla 800
 Db 2458 AAGCAAGCGGTGCAAAAGCC 2478

RESULT 2

US-09-291-922-7
 ; Sequence 7, Application US/09291922
 ; Patent No. 6383776
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Hitz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB-1163
 ; CURRENT APPLICATION NUMBER: US/09/291.922
 ; CURRENT FILING DATE: 1999-04-14
 ; EARLIER APPLICATION NUMBER: 60/083.044
 ; EARLIER FILING DATE: April 24, 1998
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 7
 ; LENGTH: 2601
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 US-09-291-922-7

Alignment Scores:

Pred. No.: 1.7e-272 Length: 2601
 Score: 2674.00 Matches: 522
 Percent Similarity: 82.20% Conservative: 92
 Best Local Similarity: 69.88% Mismatches: 115
 Query Match: 65.62% Indels: 18
 DB: 3 Gaps: 8

US-10-051-909-32 (1-800) x US-09-291-922-7 (1-2601)

QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
 Db 175 ATGAAGGTGGCGCTTGTGCTATTGCGCTTCCATTCCTTAATTTCTCCCAAGGATGG 234

QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
 Db 235 GATAATGCTACCATCGCGGGCTTAATGGTTACATTAAGAAGACCTTGTCTTTGGGAACA 294
 QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
 Db 295 -----ACTATGGAAGAGCTTGTGTGGGATGTCCTGATTTGGAGCAACGGTAATCACC 348
 QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
 Db 349 ACATGCTCTGTCTATAGCGGATTTGGCTCGGTGGCGACCCATGATGATAATCTCATCT 408
 QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu 155
 Db 409 GTGCTCTATTTCTTGGGTGGTTTGGTGATGCTGTGGTCCCCCAATGATGTATGTGTGTGC 468
 QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
 Db 469 TTGGCAGGCTACTTGTATGGATTGGGATTGGCTTGTCTGTGACTCTTGTCCCGGTCTAT 528
 QY 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
 Db 529 ATATCTGAACCGCGCGCTCTGAATTAAGGGGTGCTTGAATACGCTTCTCTCAGTTCAGT 588
 QY 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
 Db 589 GGTCTCGAGGAATGTTTGTGCTACTGTATGGTTTTCGTCATGTCATTAAGTTCGCGCG 648
 QY 216 ProAspThrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
 Db 649 CCTAGCTGGAGGCTCATGCTTGGGGTTCTGTCTATTCCTTCTCTGTATTTTGATTTG 708
 QY 236 ThrIlePheTyrLeuProGluSerProArgTyrPheLeuValSerLysGlyArgMetAlaGlu 255
 Db 709 ACCATTTTCTTGGCCGAGTCTCTCGGTGGTGTGTCAGCAAGGAAGAGTGTCTGAG 768
 QY 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeu 275
 Db 769 GCTAAGAAGGTGCTCAAGATTGCGCGAAGGAGGATGTGTGAGCGAGATGGGATTTG 828
 QY 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleLeuGlyPro 295
 Db 829 CTGGTTGAAGTCTCGGGATTGGGGTGATACATCATCGAAGAGTACATAATTTGSCCT 888
 QY 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315
 Db 889 GCTGACGATGTGGCTGATGCTCATGAACATGCAACAGAGAAGATAAAATTCGATTATAT 948
 QY 316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335
 Db 949 GGAATCCCAAGCAGGCTTCTTGGTTATCAAAACCTGTCTACTGACAGAGTTCTATTGGC 1008
 QY 336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355
 Db 1009 -----CTTGGCTCACCATGGAAGCATCATCAACCAAGCATGCCCCCTCATG 1056
 QY 356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly----- 373
 Db 1057 CATCTCTGGTGACACTGTTTGTAGCATTCATGAGAGCTCCCGGAGACAGAGCAAGA 1116
 QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
 Db 1117 GGAAGCATGCGAGCACTCTGTTTCCAAATTTTGGAAGCATGTTTCAGCACTGTGTAGCGG 1176
 QY 394 HisAlaLysAsnGluGlnTyrAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413
 Db 1177 CATGCTAAATTTGAACATGGGATGAAGAAGCTTACAAAGGAACGCTGAGGACTACATG 1236
 QY 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
 Db 1237 TCAGATCAACCCCTGGGAGCTCCGATGATAAATTTGCACAGTCCCTTAATCTCACGCCAA 1296
 QY 434 AlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGlySerAlaLeu--- 452


```
Db 522 GGCACAGATGGAAGAAGAGAGTGGCTTCAAGAGAATATATTACACCAAGATGGTGT 581
Qy 507 ProGlySerArgGlySerIleValSerLeuProGlyGlyValPheGluGly 526
Db 582 TCTGGATCTAGACGTGGTCTGTGTTTCACTCCCT--GGCGGTGATTTACCAACTGAC 638
Qy 527 SerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerIysGlyLeu 546
Db 639 AGTGAGGTGTACAGGCTGTCTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 698
Qy 547 AlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGly 566
Db 699 ATGGGTCAAGCGCAGTGGACAGCTATGATTCATCCCTCTGAACAATTCGAAAGGG 758
Qy 567 SerArgTrpIysAspLeuPheGluProGlyValArgAlaLeuValGlyValGly 586
Db 759 CCAAGTTGGAGTATCTTTTGAACCTGGGTGAAGCATGATGATTTGGTGGGTGGGA 818
Qy 587 IleGlnIleGluGlnPheAlaGlyIleAsnGlyValLeuTyrThrProGlnIle 606
Db 819 ATGCAATCTTCAGCAGTTCCTGTGTATAAATGGGTCTCTACTATAGCCCTCAAT 878
Qy 607 LeuGluGlnAlaGlyValAlaValIleLeuSerIysPheGlyLeuSerAlaSerAla 626
Db 879 CTTGAGCAGCAGGTGTGTTATCTTTCAAGCCTAGGCCTTGGTTCTACTTCTTCA 938
Qy 627 SerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet 646
Db 939 TCCTTTCTTATTAGTGGGTGACAACTTGTGATGCTTCTTGTATAGCCATTGGCATT 998
Qy 647 LeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeu 666
Db 999 AGGCTCATGATATTCAGCAGCAGAGGACTTGTCTCTAGTACATCCCGTCTTAATA 1058
Qy 667 AlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAla 686
Db 1059 GCAGCTCTTCTCATATTAGTCTGGGAAGTCTTGTGGATTTGGGATCCACTGCAATGCA 1118
Qy 687 LeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyPro 706
Db 1119 TCAATCTCAACCATAGTATTGTCTATTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1178
Qy 707 IleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAla 726
Db 1179 ATTCTTAATATATTGTGACAGATCTTCCCACTCGAGTCTGTGTCTCTGTCAATGCT 1238
Qy 727 IleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMet 746
Db 1239 ATTTGTGCCCCCTTACCTTTTGGATCTGTGATATCATTCACCTACACACTCCCGATTATG 1298
Qy 747 LeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValAlaCysLeuIleSer 766
Db 1299 CTCATTTCTGTAGGCTCGTGTGTTTGGTATTATGCTGTCTGTCTGTCTTCAATAGCA 1358
Qy 767 PheValPheValPheLeuLysValProGluThrIysGlyMetProLeuGluValIleThr 786
Db 1359 TGGGTGTGTGTCTTTTGAAGTTCAGAAACCAAGGCGCATGCCACTGGAAGTATCAT 1418
Qy 787 GluPhePheAlaValGlyAlaLysGln 795
Db 1419 GAGTTCTCTCTGTGGAGCAAAACAG 1445

RESULT 4
US-09-291-922-13
; Sequence 13, Application US/09291922
; Patent No. 638376
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
```

```
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1487
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-291-922-13

Alignment Scores:
Pred. No.: 1598-140 Length: 1487
Score: 1424.00 Matches: 281
Percent Similarity: 89.97% Conservative: 24
Best Local Similarity: 82.89% Mismatches: 32
Query Match: 34.94% Indels: 2
DB: 3 Gaps: 2

US-10-051-909-32 (1-800) x US-09-291-922-13 (1-1487)

Qy 462 GluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpIys 481
Db 12 CAGGGTGGGAGGAGTCAAGCAGCACTGGTATTGGTGGGGGGTGGCAACTCGCATGGAA 71
Qy 482 TrpSerGluLysGluGlyAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeu 501
Db 72 TGGTGGGAGCGACAGGCGGATGGCAAGAGGAGGAGGCTTCAAGAGATCTACTTG 131
Qy 502 HisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly 521
Db 132 CACCAAGAGGGGTGGCCGACTCAAGAAGGGGCTGTGTGTGTTCATCTTCTGGTGGGG 191
Qy 522 AspValPheGlu--GlySerGluPheValHisAlaAlaLeuValSerGlnSerAla 540
Db 192 GATGCCACAGAGGGGCGAGTGGTGTATACATGCTGTCTTGGTAAGCCACTCGGCT 251
Qy 541 LeuPheSerLysGlyLeuAlaGluProArgMet---SerAspAlaAlaMetValHisPro 559
Db 252 CTTTACTCAAGATCTTATGAGAGCGGTATGGGGCGCGGTCCAGCCATGATTCATCCA 311
Qy 560 SerGluValAlaAlaLysGlySerArgTrpIysAspLeuPheGluProGlyValArg 579
Db 312 TTGGAGGCGAGCTCCCAAGGTTCAATCTGGAAGATCTGTITGAACCTGGTGTAGGGGT 371
Qy 580 AlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPheAlaGlyIleAsnGlyVal 599
Db 372 GCATTTGCTCGTGTGTGGATTCAGATCTTCAGCAGTTTGTGGATTAATAGAGTT 431
Qy 600 LeuTyrTyrThrProGlnIleLeuGluAlaGlyValAlaValIleLeuSerIysPhe 619
Db 432 CTCTACTATCTCTCAAAATCTGGAGCAAGCTGGTGTGGTGTCTTCTTTTCCAATCTT 491
Qy 620 GlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeu 639
Db 492 GGCTCAGTTTCAGATCAGCATCCATCTTGATCAGTTCTCTCCACCACTTACTCATGCTC 551
Qy 640 ProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeu 659
Db 552 CCAAGCATTTGGTGTAGCCATGAGACTTATGGATATATCTGGAAGAGGTTTCTGTACTG 611
Qy 660 GlyThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAsp 679
Db 612 GGCACAAATTCCTTGTATGATCATCCCTTAATTTGGTGTGGTGTGGTGTATATCAAC 671
Qy 680 LeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCys 699
Db 672 TTGAGTACGGTGGCCCAAGCTGTGTCTCCAGATAGGCTCATTTGCTTACTTCTGCTGC 731
Qy 700 PheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArg 719
Db 732 TTTGTCTATGGCTTTGGCCCGATCCCAACATCTTATGTGTCAGAGATTTTCCCAACCA 791
```


556 MetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe----- 573
 986 -----TGGAAAGAGCTCTCTCTCTAT 1006
 574 --GluProGlyValArgAlaLeuLeuValGlyValGlyLeuGlnGln 592
 1007 CCAAGCCGCAATGTCATCATGTAACGTCCTTGGTATTCACTTCTCCACAA 1066
 593 PheAlaGlyLeuAsnGlyValLeuTyrTrpProGlnLeuLeuGlnAlaGlyVal 612
 1067 GCGTCGGCGGTACAGCCGCGTGTGTTGTACAGCCCGAGGATCTTCGAAAGGCTGGGATT 1126
 613 AlaValLeuLeuSerLysPheGlyLeuSerSerAlaSerAlaSerLeuLeuLeuSer 632
 1127 ACAACGACACGATAGCTT-----CTTGACACCGTGGCGGTGGATT 1171
 633 LeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSer 652
 1172 GTTAAGACCGTGTCTATCTTGGCG-----GCTACGTTTACGTTGGACCGCGTG 1219
 653 GlyArgArgPheLeuLeuGlyThrIleProIleLeuLeuLeuLeuValIleLeu 672
 1220 GGTGCGTCCGTTGTATTGTCTAGTGTGGCGGATGCTCTCGCTTCTCAGCTT 1279
 673 ValValSer---AsnLeuLeuAspLeuGly-----ThrLeuAlaHisAlaLeuLeuSer 689
 1280 GCGATCAGCTCTACTGTTATGATCATCGAGAGGAAATTAATGTGGCGCGTTGGATCG 1339
 690 ThrValSer---ValIleValTyrPheCysCysPheValMetGlyPheGlyProIlePro 708
 1340 AGCATAGCCATGTTGGTGTACGTGGCCACGTTCTCCATCGGTCCGCTCCCATCAG 1399
 709 AsnLeuLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCys 728
 1400 TGGGTCTATAGTCTCTGAGATCTCCGTTGAGCTCGCGCGCGCARGGTCCGCGCGGA 1459
 729 AlaPheThrPheTrpIleGlyAspIleValThrTyrSerLeuProValMetLeuAsn 748
 1460 GTTCGGGTGAATAGGACCACTAGCGGGTGTCTCAATGACTTTTCTCTCCTCCTACA 1519
 749 AlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValAlaCysLeuIleSerPheVal 768
 1520 GCATCACTATTGGTGGAGCTTCTCTTATTTGTTGGCATGCTACTGTGGTGGATA 1579
 769 PheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThr---Glu 787
 1580 TTTCTTTTACACCGTCTTCTGCTGAGACCGCGGGAACCGCTCGAAGACATGGAAGGCT 1639
 788 PhePheAlaValGlyAlaLysGlnAlaAlaAlaLysAla 800
 1640 TTTGGTACTTTTAGTCCAAATCAACCGCCAGCAGGCT 1678

RESULT 8

US-09-291-922-21
 ; Sequence 21, Application US/09291922
 ; Patent No. 6383776
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Hitz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB-1163
 ; CURRENT APPLICATION NUMBER: US/09/291,922
 ; CURRENT FILING DATE: 1999-04-14
 ; EARLIER APPLICATION NUMBER: 60/083,044
 ; EARLIER FILING DATE: April 24, 1998
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 21
 ; LENGTH: 2017
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa

US-09-291-922-21
 Alignment Scores:
 Pred. No.: 7,72e-42 Length: 2017
 Score: 494.50 Matches: 185
 Percent Similarity: 35.71% Conservative: 101
 Best Local Similarity: 23.10% Mismatches: 309
 Query Match: 12.13% Indels: 22
 DB: 3 Gaps: 22
 US-10-051-909-32 (1-800) x US-09-291-922-21 (1-2017)
 QY 24 LeuProSerValValLeuAlaLeuProGlyProLeuProProLaSerCysSerSerGln 43
 DB 30 TTACACTCGACCGCCACTACTACTATACACGGCCCA---GAGCGAGCCTCTCTCTCTGCA 86
 QY 44 GluProValThrSerAspAspIleLeuGluAsp-----LysMetSerGly--- 58
 DB 87 CCACCGGAGATGGTTCGCGCGCGCTCCGAGGCGCTCGCGCGAAGAGAGGGGCAAC 146
 QY 59 -----AlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
 DB 147 GTCCGGTTCGCTTCGCTCGCCATCTCTCGCTCCATGACCTCCATCTCTCTCTCGGCTAC 206
 QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
 DB 207 GATATCGGGGTGATGAGCGGGGCGCTGCTATCATCAAGAGGACTTCAACATC---AGT 263
 QY 96 GluProThrValGlu-----GlyLeuIleValSerMetSerLeuIleGlyAlaThr 112
 DB 264 GACGGGAGGTGGAGGTCTCATGGGCTACTGACCTCTACTCTGCTCATCGCTCTCTCTC 323
 QY 113 IleValThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIle 132
 DB 324 GCG-----GCGGGCGGACGCTCGACTGGATCGCGCGCGGTACACATCGTG 371
 QY 133 LeuSerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyr 152
 DB 372 TTCGCCCGCGTCATATTCTTCGCGGGGGSGTTCCTCATGGGTTCGCCGTCACTAGCGCC 431
 QY 153 ValLeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuVal 172
 DB 432 ATGTCATGTCGCGCGCTTCGTGGCGGCATCGCGCTGGGTACGGCTCATGCGCTCATGCGC 491
 QY 173 ProLeuTyrIleSerGluIleAlaProSerGluLeuArgGlyLeuLeuAsnThrLeuPro 192
 DB 492 CCGGTGTACACCGCGGAGGTTCGCGCGCGTGGCGCGGTCTCTGACGCTGCTCTCCG 551
 QY 193 GlnPheSerGlySerGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeu 212
 DB 552 GAGGTGTTCATCACTTCGCGCATCTGCTCGGTACGTCGTAACATATGTTTCTCTCCGC 611
 QY 213 SerPro---SerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhe 231
 DB 612 TTGCCGCTGAACTTCGCGGTGCGCATCATGCTCGCATCGCGCGCGCGCGCTCCGCTG 668
 QY 232 PhePheGlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGly 251
 DB 669 CTGCTCGGCTCATGTTGCTCGGCATCCCGAGTTCGCGCGGTGGTGGTGGTGGTGGTGG 728
 QY 252 ArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGly 271
 DB 729 CGCTCGCGGACGCAAGGTGCTCTGAGAGACCC-----TCCGACAGC 764
 QY 272 GluLeuSerLeuLeuLeuGluGlyLeuGluValGlyAspThrSerIleGluGluTyr 291
 DB 765 -----TCCGACAGC----- 773
 QY 292 IleIleGlyProAlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGln 311
 DB 774 -----GCGGAGGAGCGCGGAGCGCTG----- 797
 QY 312 IleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyPro 331

Db 797 ----- 797
 Qy 332 IleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSer 351
 Db 797 ----- 797
 Qy 352 ValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGln 371
 Db 797 ----- 797
 Qy 372 AlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThr 391
 Db 797 ----- 797
 Qy 392 AspGlnHisAlaLysAsnGluGlnTrpAspGluAsnLeuHisArgAspAspGluGlu 411
 Db 797 ----- 797
 Qy 412 TyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSer 431
 Db 797 ----- 797
 Qy 432 ArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAla 451
 Db 797 ----- 797
 Qy 452 LeuSerMetArgArgGlnThrLeuLeuGlyGlyAspGlyValSerSerThrAsp 471
 Db 798 -----GCCGAC 803
 Qy 472 IleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArg 491
 Db 804 ATCAAGCGCGCC----- 815
 Qy 492 LysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArg 511
 Db 816 -----GCCGCACTCCCTGAGGAGCTCGAC 839
 Qy 512 GlySerIleValSerLeuPro-----GlyGlyGlyAspValPheGluGlySerGluPhe 529
 Db 840 GCGCAGCTGTGACCGTCCCAAGAGAGGAGCGGAAAC----- 878
 Qy 530 ValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluPro 549
 Db 879 -----GAGAAG 884
 Qy 550 ArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrp 569
 Db 885 CGGGTG-----TGG 893
 Qy 570 LysAspLeuPhe-----GluProGlyValArgAlaLeuLeuValGlyValGly 586
 Db 894 AAGGAGCTCATCTGTCGCCGACCCCGCATCGCGCATCTGCTGTCCGGGATCGGC 953
 Qy 587 IleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTrpProGlnIle 606
 Db 954 ATCCACTTCTCCAGCATGGTGGGATTCACCTCGCTGTCTTACACCCCTCTCGG 1013
 Qy 607 LeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAla 626
 Db 1014 TTCAAGAGCCCGGATTAACG-----AACGACAAACACTTCTTGGGCACCACTTGGCCG 1067
 Qy 627 SerIleLeuLeuSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet 646
 Db 1068 TTCGGTGTCCCAAGAGCGCTTTTCATCTTTG-----GCCACT 1106
 Qy 647 LeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIle 666
 Db 1107 TTCTTCATCAGCGCTCGCGCGCGCGCTGTGTCTGGGAGCAGCGCGGGGATAATC 1166
 Qy 667 AlaSerLeuVal-----IleLeuValValSerAsnLeuIleAspLeuGly 681

Db 1167 CTCCTCCTCATCGCGCTCGCGCGCTCACCGTGTGCGCCAGCACCCCGAGCCCAAG 1226
 Qy 682 ThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheVal 701
 Db 1227 ATACCTCGGGCCATCGCGCTAAGCATCGCCTCCACCCCTCGCTACGTGCGCTTCTCTCC 1286
 Qy 702 MetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArg 721
 Db 1287 ATCGGCTTGGCCCATCATCGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGGC 1346
 Qy 722 -----GlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleVal 739
 Db 1347 GCGCTGGCGTGTGCTCGCGCTCGCGCCAAACCGCGTCACACGCGGTCTCTCTCATG 1406
 Qy 740 ThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyr 759
 Db 1407 ACCCTCTG-----TCGCTGTCCAGGCAATCACCACCGCGGAGCTTCTCTCTAC 1460
 Qy 760 AlaValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGly 779
 Db 1461 TCGGCACTCGCGCTCGCTCGCGTGTCTTCTACACCTACCTCCCGGAGACCGCGGC 1520
 Qy 780 MetProLeuGluValIleThrGluPheAlaValGlyAlaLysGlnAlaAlaLys 799
 Db 1521 CGGACGCTGGAGGAGATGAGCAAGCTTTC-----GCCGACAGCGCGCGCTCGGAA 1574
 Qy 800 Ala 800
 Db 1575 TCA 1577
 RESULT 9
 US-09-291-922-19
 ; Sequence 19, Application US/09291922
 ; Patent No. 6383776
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Hitz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB-1163
 ; CURRENT APPLICATION NUMBER: US/09/291,922
 ; CURRENT FILING DATE: 1999-04-14
 ; EARLIER APPLICATION NUMBER: 60/083,044
 ; EARLIER FILING DATE: April 24, 1998
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 19
 ; LENGTH: 1914
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; US-09-291-922-19
 Alignment Scores:
 Pred. No.: 1,46e-41 Length: 1914
 Score: 491.50 Matches: 172
 Percent Similarity: 35.36% Conservative: 107
 Best Local Similarity: 21.80% Mismatches: 201
 Query Match: 12.06% Indels: 309
 DB: 3 Gaps: 18
 US-10-051-909-32 (1-800) x US-09-291-922-19 (1-1914)
 Qy 41 SerSerGlnGluProValThr-----SerAspAspIle-----LeuGluAspLys 55
 Db 35 TCAAGAGAGTACCGCTTAACGATGGCTTCCGACGAGCTCGCAAGCGGTGAGGCCAGG 94
 Qy 56 MetSerGly-----AlaValLeuValAlaIleValAlaSerIleGlyAsnLeu 71
 Db 95 AAGAGAGGCAACGTCAGTATGCTTCCATATGTGCACTCTGCGCTCCATGCGCTCTGTC 154
 Qy 72 LeuGlnGlyTrpAspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPhe 91

D	155	ATCCTTGGCTATGACATGGGTGATGAGTGGAGCGGCCATGTACATCAAGAGGACCTG	214	Q	450	SerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyAspGlyValSerSer	469
Q	92	GlnLeuGlnAsnGluProThrValGluGlyLeuLeuValSerMetSerLeuLeuGlyAla	111	D	760	-----	760
D	215	AATATC---ACGGACGTGACGTGGAGATCCTGATCGGATCTCATCTCTAC-----	265	Q	470	ThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsn	489
Q	112	ThrIleValThrThrPheSerGlyPro---LeuSerAspSerIleGlyArgArgProMet	130	D	761	---GACATCAAGCGCGG-----	775
D	266	TGCTGTTCGGATCCTTCGTGGCGCGCGGACGTCGACAGGATCGGCGCGCTTGACC	325	Q	490	GlyArgLysGluGlyGlyPheLysArgValTrpLeuHisGlnGluGlyValProGlySer	509
Q	131	LeuIleLeuSerSerIleLeuTrpPhePheSerGlyLeuIleMetLeuTrpSerProAsn	150	D	776	-----GCGGGGATTCCGAAGGCG	793
D	326	GTGCTGTTCGCGCTGTCATCTTCGTGGGCTCGTGTCTCATGGGTTTCGCGCTCAAC	385	Q	510	ArgArgGlySerIleValSerLeuPro-----GlyGlyGlyAspValPheGlu	525
Q	151	ValTrpValLeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThr	170	D	794	CTCAGCGGGGACGTAGTACCGGTACCGCGGAGAGGAGGCGCGGTGAGTTG-----	847
D	386	TACGGCATGTCTATGGCGGCGCGCTTCGTGGCGGAGTCGTGTGGCTACGGGGGCATG	445	Q	526	GlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGly	545
Q	171	LeuValProLeuTrpIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThr	190	D	847	-----	847
D	446	ATCCGCGCGGTGTACACGGCGGAGATCTCCCTGGCGGCTCCGCTTCTTGACACACC	505	Q	546	LeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLys	565
Q	191	LeuProGlnPheSerGlySerGlyMetPheLeuSerTrpCysMetValPheGlyMet	210	D	847	-----	847
D	506	TTCCCGGAGGTTCATCAACATCGGCATCTCTGCTTGGCTACCTGTCCAACTTCGCGTTC	565	Q	566	GlySerArgTrpLysAspLeuPhe-----GluProGlyValArgAlaLeuLeu	582
Q	211	SerLeuSerPro---SerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSer	229	D	848	---CAGGTGTGGAAGAGCTCATCTGTCCCGACCCCGGTGTCCGACGATACTGCTC	904
D	566	CGCGGCTCCGCTCCACCTCGGCTGGCGGTGATGCTGCCATTCGGGAGTTCGCTCC	625	Q	593	ValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTrpTyr	602
Q	230	LeuPhePhePheGlyLeuThrIlePheTrpLeuProGluSerProArgTrpLeuValSer	249	D	905	TCGCGCCTGGTCTCCACTTCTCCAGCAGCTTCTGCGACGCACTCCGTCGTCAGTAC	964
D	626	---GGCCTGCTCGGCTCTCTGTTCTGCATCCCGAGTCGCTCGGTGGTGGTCTTG	682	Q	603	ThrProGlnIleLeuGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSer	622
Q	250	LysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspVal	269	D	965	AGCGCCGCTGTTCAAGCGCGGGATCACCGACACCAACAGCTCTCTGGGCGTCACC	1024
D	683	AAGGGCGCTCGCGGACCGCGGTGCTAGAGAAG-----	721	Q	623	SerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIle	642
Q	270	SerGlyLeuLeuSerLeuLeuGluGlyLeuGluValGlyAspThrSerIleGlu	289	D	1025	TGCGCG-----GTGGGCTGACCAAGACGTTCTTCATCTG-----	1060
D	721	-----	721	Q	643	GlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIle	662
Q	290	GluTrpIleIleGlyProAlaThrGluAlaAspAspLeuValThrAspGlyAspLys	309	D	1061	---GTGGCCACGTCTCTGACCGCGCGGGGCTCGGCTCTGCTCTGATCAGCACG	1117
D	721	-----	721	Q	663	ProIleLeuIleAlaSerLeuValIleLeu-----ValValSerAsnLeu	677
Q	310	GluGlnIleThrLeuTrpGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLys	329	D	1118	GGCGGGATGATTGCTCGCTCATCTGCTCGGGTCGGGGCTCACCGTCGGGGGCATCAC	1177
D	722	---ACCTCTGCCACGCCAGAGGAGCGCGGAGCGGCTGCGC-----	760	Q	678	IleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTrpPhe	697
Q	330	GlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsn	349	D	1178	CCGACACCAAGTCGCGTGGCGCGCTCGCCTGTGCATCGCTCAACCCCTGCTCTATC	1237
D	760	-----	760	Q	698	CysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePhePro	717
Q	350	GlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMet	369	D	1238	GCTTCTTCTCCATCGGCTCGGCGCCATCACGCGCGCTGTACACCTCGGAATATTCCCG	1297
D	760	-----	760	Q	718	ThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIle	737
Q	370	ProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSer	389	D	1298	CTGAGGTGCGCGCTGGGCTTCGCGTGGGTGTGGCGAGCAACCCGCTCACCGAGCGCC	1357
D	760	-----	760	Q	738	IleValThrTrpSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSer	757
Q	390	ValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAsp	409	D	1358	GTCACTCCATGACTCTCTGCTCCCAAGGCCATCACCATCGCGCGGAGCTTCTTC	1417
D	760	-----	760	Q	758	IleTrpAlaValValCysLeuIleSerPheValPheValPheLeuLysValProGluThr	777
Q	410	GluGluTrpAlaSerAspGlyAlaGlyGlyAspTrpGluAspAsnLeuHisSerProLeu	429	D	1418	CTCTACTCGGCATCGCGCGCTGCTGGGTTCCTTCTTCTTCTTCTTCTTCTTCTTCTT	1477
D	760	-----	760	Q	778	LysGlyMetProLeuGluValIleThrGluPhePhe-----Ala	790
Q	430	LeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGly	449	D	1478	CGCGCGGAGCTGTGAGAGATGGGCAAGCTGTTGCGCATGCCAGACACGGGCGATGGCT	1537
D	760	-----	760				


```
Db 920 -----|||...|||
; TGAAGGAGCTCATCTTTTCGCCGACCCAGCCATG 955
Qy 578 ArgArgAlaLeuLeuValGlyValGlyLeuLeuGlnPheAlaGlyLeuAen 597
; |||...|||
Db 956 CGCGCATATGCTCGCGGCTCGGCATCCATTTCTTTCAGCAGCGGCGCTCCGAC 1015
; |||...|||
Qy 598 GlyValLeuTyThrProGlnLeuLeuGlnAlaGlyValAlaValLeuSer 617
; |||...|||
Db 1016 TCGCTGCTGTATAGCCACGCGTGTCCAGAGCGCGGCATCACCGCGCACCAACCAC 1075
; |||...|||
Qy 618 LysPheGlyLeuSerAlaSerAlaSerIleLeuLeuSerLeuThrThrLeuLeu 637
; |||...|||
Db 1076 CTGCTCGC-----GCCATGCGCCATGGGGTCAAGAAGCGCTCTTC 1120
; |||...|||
Qy 638 MetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeu 657
; |||...|||
Db 1121 ATCTG-----GTGGCAGCTTCCAGCTCGACCGCGTCCGACAGGCGCGCTG 1168
; |||...|||
Qy 658 LeuLeuGlyThrIleProIleLeuLeuAlaSerLeuVal-----IleLeu 672
; |||...|||
Db 1169 CTGCTGACCGACGCGCGCATGCTCGCTGTCTCATCGGCCCTCGGACGGGCTCAC 1228
; |||...|||
Qy 673 ValValSerAsnLeuLeuAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSer 692
; |||...|||
Db 1229 GTCGTGGTGGCCACCGGACGCAAGTCCCGTGGGCCATCGGCTGTGTCATCGTGTCC 1288
; |||...|||
Qy 693 ValIleValTyPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys 712
; |||...|||
Db 1289 ATCTGCGCTACGTCCTCTCTCTCATCGGCTCGGCGCCCTCACCGCGGTACAC 1348
; |||...|||
Qy 713 AlaGlyLeuPheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPhe 732
; |||...|||
Db 1349 TCGAGGTCTTCCACTCGGGTGGCGGCTGGCTTCGCGTGGGACGTCATGCAAC 1408
; |||...|||
Qy 733 TriPleGlyAspIleLeuValThrTyThrSerLeuProValMetLeuAsnAlaIleGlyLeu 752
; |||...|||
Db 1409 CGGCTCACCGACGCGCGGTCTCCATGCTCTCTCTCTGCTTGTCCAAAGCCATCAC 1468
; |||...|||
Qy 753 AlaGlyValPheSerIleTyAlaValValCysLeuLeuSerPheValPheValPheLeu 772
; |||...|||
Db 1469 GCGCGAGCTTCTTCGTACGCGGCGCATCGCGCGATAGGATGATTTCTTCTCAC 1528
; |||...|||
Qy 773 LysValProGlnTyThrLysGlyMetProLeuGluValIleThrGluPhePhe----- 789
; |||...|||
Db 1529 TTCATTCCGAGACGCGTGGCTCGCGCTCGAGAGATAGGAAGCTTTTCGCGCATGACG 1588
; |||...|||
Qy 790 -----AlaValGlyAlaLysGlnAlaAlaLys 799
; |||...|||
Db 1589 GACACGCGCTCGAAGCCCAAGACACCGCCACCAAA 1624
```

RESULT 11

```
US-09-291-922-25
; Sequence 25, Application US/09231922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 25
; LENGTH: 2089
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-291-922-25
```

```
Alignment Scores: 1.58e-38 Length: 2089
Pred. No.: 163.50 Matches: 166
Score: 35.82% Conservative: 108
Percent Similarity: 21.70% Mismatches: 200
Best Local Similarity: 11.37% Indels: 291
Query Match: 3 Gaps: 17
DB: 3

US-10-051-909-32 (1-800) x US-09-291-922-25 (1-2089)

Qy 52 LeuGluAspLysMetSerGly-----AlaValLeuValAlaIleValAlaSer 67
; |||...|||
Db 104 GTCGAGCCCAAGAGAGAGGCAAGCTGAGGTTCGCTTCGCTGCGCATCTCTCGCTCC 163
; |||...|||
Qy 68 IleGlyAsnLeuLeuGlnGlyTyPheAspAsnAlaThrIleAlaAlaValLeuTyIle 87
; |||...|||
Db 164 ATGACCTCCATCTCTCTCGCTAGACATCGGCGTGTATGAGCGGAGCGTCTGCTGATC 223
; |||...|||
Qy 88 LysLysGluPheGlnLeuGlnAsnGluProThrValGlu-----GlyLeuIleVal 104
; |||...|||
Db 224 CAGAGGATCTGAGATC---NAGCACCCAGCTGAGGTCTCTCATGCGCATCTCTCAAC 280
; |||...|||
Qy 105 SerMetSerIleIleGlyAlaThrIleValThrThrPheSerGlyProLeuSerAspSer 124
; |||...|||
Db 281 GTGTACTCGCTCATTTGCTCTCTCGC-----GCGGCGGAGCTCGGACTGG 328
; |||...|||
Qy 125 IleGlyArgArgProMetLeuLeuLeuSerSerIleLeuTyThrPhePheSerGlyLeuIle 144
; |||...|||
Db 329 ATCGGCGCGGCTTCACATCGCTTCGCGCGGTGATCTTCTTCGCGGCGGCGCTCATC 388
; |||...|||
Qy 145 MetLeuTrpSerProAsnValTyValLeuLeuLeuAlaArgPheValAspGlyPheGly 164
; |||...|||
Db 389 ATGGGCTTCTCGTCAACTACGCCATGCTCATGTTCGGGCGCTTCGTGGCGCGCATCGC 448
; |||...|||
Qy 165 IleGlyLeuAlaValThrLeuValProLeuTyIleSerGluIleAlaProSerGluIle 184
; |||...|||
Db 449 GTGGGTACGCTCTCATGATCGCGCGCTGAACACGCGGAGGTGTCCTCCGCTGTGCC 508
; |||...|||
Qy 185 ArgGlyLeuLeuAenThrLeuProGlnPheSerGlySerGlyGlyMetPheLeuSerTy 204
; |||...|||
Db 509 CGTGGGTCTCATACCTTCCGAGAGGTGTTCACTCAACTTCGCACTCTCTCGGATAT 568
; |||...|||
Qy 205 CysMetValPheGlyMetSer---LeuSerProSerProAspTrpArgIleMetLeuGly 223
; |||...|||
Db 569 GTCTCCAACTTCGCTTCGCGCGCTCTCCCTCGCGCTCGGCGCATATGCTCGGC 628
; |||...|||
Qy 224 ValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyLeuProGluSer 243
; |||...|||
Db 629 ATAGCGCGGTGCGTCCGTC---CTGCTCGGTTCATGGTCTCGCATGCCGAGTCT 685
; |||...|||
Qy 244 ProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysValLeuGlnLysLeu 263
; |||...|||
Db 686 CCGCGTGGTCTCATGAAGGCGCGTCTCGGAGCGCAAGGTGTGTCTGCGCAAGACG 745
; |||...|||
Qy 264 ArgGlyLysAspValSerGlyGluLeuSerLeuLeuGluGlyLeuGluValGly 283
; |||...|||
Db 745 ----- 745
; |||...|||
Qy 284 GlyAspThrSerIleGluGlyTyIleIleGlyProAlaThrGluAlaAlaAspLeu 303
; |||...|||
Db 746 TCCGACACG----- 754
; |||...|||
Qy 304 ValThrAspGlyAspLysGluGlnIleThrLeuTyGlyProGluGluGlyGlnSerTip 323
; |||...|||
Db 755 -----CCGGAAGAGCGCGCGGCGCGC 775
; |||...|||
Qy 324 IleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArg 343
; |||...|||
Db 776 ATCGCC----- 781
; |||...|||
Qy 344 HisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGly 363
; |||...|||
```

Db 781 ----- 781
 Qy 364 SerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsn 383
 Db 781 ----- 781
 Qy 384 PheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluInTrpAspGluGlu 403
 Db 781 ----- 781
 Qy 404 AsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAsp 423
 Db 782 ----- 799
 Qy 424 AsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleVal 443
 Db 799 ----- 799
 Qy 444 HisHisGlyHisArgGlySerAlaLeuSerMetArgGlnThrLeuLeuGlyGluGly 463
 Db 799 ----- 799
 Qy 464 GlyAspGlyValSerSerThrAspIleGlyGlyTyrGlnLeuAlaTrpLysTrpSer 483
 Db 799 ----- 799
 Qy 484 GluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGln 503
 Db 799 ----- 799
 Qy 504 GluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspVal 523
 Db 800 ---GGCATCCCTCTGGCCCTCGAGCGGACGTGTCGCCGTGCCAAA----- 844
 Qy 524 PheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSer 543
 Db 845 -----AAC 847
 Qy 544 LysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAla 563
 Db 848 AAAGGAAGACGAGGAGAGACGGCTTTTGAAGAGACCTCATCTGTCACGACCATAGCC 907
 Qy 564 AlalysGlySerArgTrpLysAspLeuPheGluProGlyValArgAlaLeuLeuVal 583
 Db 908 -----ATGGCCACATCCCTCATCGG 928
 Qy 584 GlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThr 603
 Db 929 GGAATCGGCATCCACTTCTTCCAGCAGTCTTCGGGCATCGACGCCCTGCTCTACAGC 988
 Qy 604 ProGlnIleLeuGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSer 623
 Db 989 CCGTAGTTTTCAGAGCGCGCGCATCAG-----GGCGACAGCGT 1030
 Qy 624 ---AlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIle 642
 Db 1031 CTCGGCGGCACACCGTGGCGGTCCGGGCCACCAATACGGTCTCATCTG----- 1081
 Qy 643 GlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIle 662
 Db 1082 ---GTGGCCACCTTCTCTCAGCCGATCCGGCGCGCGCGTGGGCTGACACGACG 1138
 Qy 663 ProfileLeuIleAlaSerLeuVal-----IleLeuValValSerAsnLeu 677
 Db 1139 GGCGGCATGCTCGTCTCTTAGTGGGCTCCGACGCGGGCTCACCGTCAACAGCGCCAC 1198
 Qy 678 IleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPhe 697
 Db 1199 CCGGACGAGAGATCAGCTGGGCCATCGTCTGTGATCTTCTGCAATCAGGCTCAGTG 1258
 Qy 698 CysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePhePro 717
 Db 1259 GCCTTCTTCCATCGCGCTCGGCCCATCAGTGGGTGTACAGCTCGGAGTCTCCCG 1318

Qy 718 ThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIle 737
 Db 1319 CTGCACGTGCGCGCGTGGCTGCTCCCTGGCGGTGGCGGTCAACCGCTGACGAGCGC 1378
 Qy 738 IleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSer 757
 Db 1379 GTGATCTCCATGACCTTCATTTCCGTCGCAAGCCATGACCATCGGCGGCGCTTCTTC 1438
 Qy 758 IleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysValProGluThr 777
 Db 1439 CTCITTCGCGCATCGCTCATTCGTCATGGTGTCTTCTTCGCTACCTCCGCGAGACC 1498
 Qy 778 LysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla-----LysGln 795
 Db 1499 CGCGCCCGACGCTGGAGCATGAGCTCGCTGTCGCAACACGCGCCACGCAAGCAG 1558
 Qy 796 AlaAlaAlaLysAla 800
 Db 1559 GCGCGCGGAGAGCC 1573
 RESULT 12
 US-09-489-039A-4762
 ; Sequence 4762, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 4762
 ; LENGTH: 1431
 ; TYPE: DNA
 ; ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-4762
 Alignment Scores:
 Pred. No.: 4,548-37 Length: 1431
 Score: 447.00 Matches: 158
 Percent Similarity: 33.68% Conservative: 98
 Best Local Similarity: 20.75% Mismatches: 186
 Query Match: 10.97% Indels: 318
 DB: 4 Gaps: 18
 US-10-051-909-32 (1-800) x US-09-489-039A-4762 (1-1431)
 Qy 42 SerGlnGluProValThrSer---AspAspIleLeuGluAsp---LysMetSerGly--- 58
 Db 4 TCACAGATCAATCACTCACTGAGGCGGTATGCTGACACAAACAAAGGCGT 63
 Qy 59 -----AlaValLeuValAlaIleValAlaSerIleGlyAsnLeuGln 73
 Db 64 TCGAACAGACATGACGTTCTGCTGTTTCTCGCGCGGTGCTGCTGCTGCTGCT 123
 Qy 74 GlyTyrAspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeu 93
 Db 124 GCGCTGTATATCGGTGTTATTCGCGGTGCTTACCTTATTCGCAATGAGTTCAGATT 183
 Qy 94 GlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIle 113
 Db 184 TCCGCCAC-----ACCAGGAGTGGTGGTTCAGTCCATGATGTTGCGGCTGCCGTC 237
 Qy 114 ValThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeu 133
 Db 238 GCGCGCGTGGCAGCGGTGCTCTCTTCAACTGGGCGGAGAAAGAGCGCTGATGTC 297
 Qy 134 SerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrVal 153

Db 298 GGGCCATCCTCTTCGTCGGCGTTCGCTGTTCTCTCGCCGCCGCAAAACGTCGAGATC 357
Qy 154 LeuLeuAlaArgPheValAspGlyPheGlyLeuAlaValThrLeuValPro 173
Db 358 CTGCTGTTTCCCGTGTCTCGGCTCGCGCTGGCGTGCCTCATATACGGCTCG 417
Qy 174 LeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGln 193
Db 418 CTGTATCTGTCGGAATCGCCCAAAAAATCGCGCAGTATGATTTCCATGTACACAG 477
Qy 194 PheSerGlySerGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213
Db 478 CTGATGATCACCATCGGATCTTGGCGCCTAT-----CTCTCTGACACCGCTTTTCAGC 531
Qy 214 ProSerProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233
Db 532 TACAGCGCGCATGGCGCTGATGCTCGGGTTATCATCTCCGGCGGTTTGTGCTG 591
Qy 234 GlyLeuThrIlePheTyrLeuProGluSerProArgTyrLeuValSerLysGlyArgMet 253
Db 592 ---ATCGCGCTTATCTCTCGCGACAGCCCGCTGTTCCGCCCAACGTCGCTTT 648
Qy 254 AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyLeu 273
Db 649 GTCGATCGGAACGGCTGCTGCTCGCTCGC----- 681
Qy 274 SerLeuLeuLeuGlyLeuGluValGlyAspThrSerIleGluGluTyrIle 293
Db 682 -----GATACCAAGCCGCA----- 696
Qy 294 GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThr 313
Db 696 ----- 696
Qy 314 LeuTyrGlyProGluGluGlnGlnSerTrpIleAlaArgProSerLysGlyProIleMet 333
Db 696 ----- 696
Qy 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
Db 696 ----- 696
Qy 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
Db 696 ----- 696
Qy 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db 696 ----- 696
Qy 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspGluGluTyrAla 413
Db 697 ---GCGAAACCGCGCTC----- 711
Qy 414 SerAspGlyAlaGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
Db 712 ---GATGAATCCGTGAAGCCGTGAAGCTGAAGCTAAACAG 744
Qy 434 AlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGlySerAlaLeuSer 453
Db 745 TCC----- 747
Qy 454 MetArgGlnThrLeuLeuGlyGluGlyAspGlyValSerSerThrAspIleGly 473
Db 747 ----- 747
Qy 474 GlyGlyTrpGlnLeuAlaTrpLysTrpSerGluGlyGluAsnGlyArgLysGlu 493
Db 748 ---GGCTGGTGGCTG----- 759
Qy 494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer 513
Db 760 -----TTTAAA----- 765

Qy 514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533
Db 766 -----GACAAACAGCACTTC----- 780
Qy 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553
Db 780 ----- 780
Qy 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTyrLysAspLeuPhe 573
Db 780 ----- 780
Qy 574 GluProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPhe 593
Db 781 -----CGCCGCGGCTGTTCTCGGCTACCTGCTGAGGTGATGCAACAGTTC 828
Qy 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613
Db 829 ACCGGGATGAACGTATCATGCTACGCGCGAAGATCTTTGAGCTGGCGGTTATGCC 888
Qy 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
Db 889 AACACCATGAGCAAAATGTGG-----GGGACAGTGATCGTCGGTCTCACT 933
Qy 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
Db 934 AACGTGCTGGCCACTTTATCGCCATCGGT-----CTGGTCGACCGCTGGGGC 981
Qy 654 ArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
Db 982 GTTAAA-----CCGACGCTGATCTTGCCTTATCGTATGGCC 1020
Qy 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHis----- 685
Db 1021 GCGGGAATGGCGTC---CTGGGTACCATGATGATCGCATTCCTCTCTACCGCC 1077
Qy 686 AlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGly 705
Db 1078 CAGTACATCGCGCTCTGATGCTGCTGATGTTTCATGCTGCTTCCCATGAGCGCGCC 1137
Qy 706 ProfileProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIle 725
Db 1138 CCACGTATTTGGGTACTGTCGGAATCCAGCCGCTGAAAGCGCGGACTTCGGTATC 1197
Qy 726 AlaIleCysAlaPheThrPheThrIleGlyAspIleValThrTyrSerLeuProVal 745
Db 1198 ACCTGCTCCACGACGCAACATGATTCGCAATGTCGCGCCACCTTCCTGACC 1257
Qy 746 MetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIle 765
Db 1258 ATGCTCAACTCGCTGGCGAGCCCAATACCTTCTGGGTGTACGCGCTCTGAACGTGCTG 1317
Qy 766 SerPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIle 785
Db 1318 TTTATCTCTGACGCTGCTGCTGATCCCGGNAACCAAAACGCTCTGCTGGAACATATT 1377

RESULT 13

US-09-489-039A-4731
; Sequence 4731, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4731
; LENGTH: 1545


```

; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-4731

Alignment Scores:
Pred. No.:      2,546-36      Length:      1545
Score:          440-50      Matches:      164
Percent Similarity: 34.05%    Conservative: 105
Best Local Similarity: 20.76%  Mismatches:      206
Query Match:      10.81%     Indels:       315
DB:              4          Gaps:         16

US-10-051-909-32 (1-800) x US-09-489-039A-4731 (1-1545)

Qy 25 ProSerValValLeuAlaLeuProGlyProLeuProProLaSer----- 39
Db 37 CCATATCTTTGGCTTATCTGCTGACCAATCATCCACGATCCAGTCTTCGTGCTTA 96
Qy 40 -----CysSerSerGlnGlu-----ProValThrSerAspSrlLeuGlu 53
Db 97 TGTCTTACTCTGTGGCAGGAAAAATGACTTCAATCAGTAACGACTCTACATTATCG 156
Qy 54 AspLysMetSerGlyAlaVal-----LeuValAlaIleValAlaSerIle 68
Db 157 CGCGGAGCAACCGTATACCCGCGGATGAACGTGTTGTTCATCGCGCGCGGGT 216
Qy 69 GlyAsnLeuGlnGlyTirAspAsnAlaThrIleAlaAlaValLeuTyrIleLys 88
Db 217 GCTGGCTTCTCTTTGGCTGGATATCCGCGTATATCCGAGGGTTGCCCTTTAACC 276
Qy 89 LysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeu 108
Db 277 GACCATTTCCACTTATCCAGCCAG-----CTTCAGGAGTGGGTGGTTAGCAGCATGATG 330
Qy 109 IleGlyAlaThrIleValThrPheSerGlyProLeuSerAspSrlLeGlyArgAg 128
Db 331 TTGGGGGCGCGGATAGGCGCGTGTTTAACGCTGGCTGCTTTCCGCTTGGCCGGA 390
Qy 129 ProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuLeuMetLeuTyrSp 148
Db 391 TACAGCCTGATGCGGGGCGCTACTCTTTGTGGCGGCTCTATCGGATCCGCTTTGGCC 450
Qy 149 ProAsnValTyrValLeuLeuLeuAlaArgPheValaspGlyPheGlyIleGlyLeuAla 168
Db 451 GCCAGCGTGAGGTGTGTGTGCGCCCGCTGGTGTGGCGGTGGCAGTCGGGATGGCC 510
Qy 169 ValThrLeuValProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeu 188
Db 511 TCTTATACCGCGCGCTGTACTCTCCAGATGCCACGAGACGTGCCGGGAAATG 570
Qy 189 AsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeuSerTyrCysMetValPhe 208
Db 571 ATCAGTATGTATCAGCTGATGTCACCCCTTGGCATTTGCTGGCGTT-----CTTTCC 624
Qy 209 GlyMetSerLeuSerProSerProAspThrPargIleValLeuAlaIlePro 228
Db 625 GATACCGCCTTTAGCTACAGCGGTAACTGGCGCGCCATCTGTGGCGGTGCTGGCGCTGCCG 684
Qy 229 SerLeuPhePheGlyLeuThrIlePheTyrLeuProGluSerProArgTirLeuVal 248
Db 685 GCGGTGATCCTGATCATTTCTGGTCTGCTTT---TTGCCGAACAGCCGCGCTGGCTGGCG 741
Qy 249 SerLysGlyArgMetAlaGluAlaLysIysValLeuGlnLysLeuArgGlyLysAspAsp 268
Db 742 GAGAAGGAGCGCCATATCGAAGCGGAAGAGTCTGGCGATGCTGCGC-----789
Qy 269 ValSerGlyGluLeuSerLeuLeuGluGlyLeuGluValGlyIlyAspThrSerIle 288
Db 790 -----GATACCTCG---798
Qy 289 GluGluTyrIleIleGlyProLathrGluAlaAlaAspAspLeuValThrAspGlyAsp 308
Db 799 -----GAAAAGGCGCGCAGAGCTT-----819

```

```

QY 668 SerLeuValIleLeuValSerAsnLeuLeuAspLeuGlyThrLeuAlaHisAlaLeu 687
Db 1135 ACITGGTGTGGCTGGCTACTGATGACATTCGAAATGGACCC---GATCCAGCGC 1191
QY 688 LeuSerThrValSerValIleValTyrPheCysCys-----PheValMetGlyPhe 704
Db 1192 CTCTCTGGCTCTCCGTCGCGCATGACCATCATGTATTGCGGGTATGCGATGACGCG 1251
QY 705 GlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCys 724
Db 1252 GCGCGGTGTGGTGGTCTCTGCTGAGATCCAGCCGCTAAATGCGCGCACTTCGGT 1311
QY 725 IleAlaIleCysAlaPheThrPheThrIleGlyAspIleValThrTyrSerLeuPro 744
Db 1312 ATCACCCTGCTCGACCAACCACTGGGTGTCGAACATGATCATCGCGCCACTTTCCTG 1371
QY 745 ValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeu 764
Db 1372 ACCTGCTTACACGCGATGGCGCGCGGCACCTTCTGCTCTACACGCGCTCAACGTG 1431
QY 765 IleSerPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluVal 784
Db 1432 GCCTTTATCGGCATCACCCTCTGGCTGATCCCGAACCAGATGTCACCTCGAGCAC 1491
QY 785 IleThrGluPhePheAlaValGlyAlaLys 794
Db 1492 ATTGAGCGCAACCTGATGGCGGCGAGAG 1521

```

RESULT 14

```

US-09-489-039A-4560
; Sequence 4560, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Berton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4560
; LENGTH: 1506
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4560

```

```

Alignment Scores:
Pred. No.: 6 52e-35 Length: 1506
Score: 427.00 Matches: 155
Percent Similarity: 33.29% Conservative: 104
Best Local Similarity: 19.92% Mismatches: 201
Query Match: 10.48% Indels: 318
DB: 4 Gaps: 17

```

US-10-051-909-32 (1-800) x US-09-489-039A-4560 (1-1506)

```

QY 33 GlyProLeuProAlaSerCysSerSerGlnGluPro-----ValThrSer 48
Db 19 GGGCCG---CCTGAACCCCTACTGTTCCCTCGCTTTCCCTTATAGAGGAATCATTAAGAC 75
QY 49 AspAspIleLeuGluAspLysMetSerGlyAlaValLeuValAlaIleValAlaSerIle 68
Db 76 AACCGCAGACACATCTGAATGGCTACGCTGGACGATTTGCTGGTCCGCCCTGC 135
QY 69 GlyAsnLeuGlnGlyTrpAspAsnAlaThrIleAlaAlaValValTyrIleLys 88
Db 136 GGTGGTTTACTGTTTGGCTATGACTGGGTGGTCAITGGCGGCGCTAAAGCCATTTATGAA 195
QY 89 LysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeu 108
Db 196 GCTGTGTTTCAATT---ACGACCCGCGCGAGTCCGCTGGCGGCTGAGCTCAGCGCTG 252

```

```

QY 109 IleGlyAlaThrIleValThrPheSerGlyProLeuSerSerIleGlyArgArg 128
Db 253 TTGGGCTGATATTTTCGGCGCATTAATTTTCGGATGGTGCACAGACAAATCGGGCGCAAG 312
QY 129 ProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSer 148
Db 313 CTGCCATTAATTTCTTCGGCGCTGCTTTCAGCGCGTGGCTGGGGACGGGGTGCCTGCC 372
QY 149 ProAsnValTyrValLeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAla 168
Db 373 AGTCATTTCGATATGTTGTGTTTACCGCATTTGTGGCGCGCTAGGGATTTGCTGGCT 432
QY 169 ValThrLeuValProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeu 188
Db 433 TCCGCTCTCAGCCCGCTTTACATTCGCGAAGTACAGCCCGCGAGAGAAAGAGGACGTTT 492
QY 189 AsnThrLeuProGlnPheSerGlySerGlyMetPheLeuSerTyrCysMetValPhe 208
Db 493 GTCGCGGTCAATCAGCTCACCATCGTATTGGCGTGTCTGCGCGCTCAGTTAATCAATCTG 552
QY 209 GlyMetSerLeuSerProSerPro----- 216
Db 553 ATGATTCTGAACCGGTGGAGCCCGGGCGAGCAGCAGATGATTGTGGACACGCTGGAAT 612
QY 217 -----AspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233
Db 613 GGGCAGATGGCTGGCGCTGATGTTTCGTCGCGAACTGGTGGCGGCACTGGCGTTCTG 672
QY 234 GlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMet 253
Db 673 GTCCTG---ATGTTTTTGTCCCGAGTCGCGCGTGGCTCATGAAGCGCGGTAAACCG 729
QY 254 AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeu 273
Db 730 GAGCGCGCGCGGTGGCTGGTGAACGC----- 756
QY 274 SerLeuLeuLeuGluGlyLeuGluValGlyAspThrSerIleGluGluTyrIleIle 293
Db 757 -----ATT 759
QY 294 GlyProAlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGlnIleThr 313
Db 760 GGTCTGCTC----- 768
QY 314 LeuTyrGlyProGluGluGlnSerTrpIleAlaArgProSerLysGlyProIleMet 333
Db 768 ----- 768
QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
Db 768 ----- 768
QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
Db 768 ----- 768
QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db 768 ----- 768
QY 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspGluGluTyrAla 413
Db 768 ----- 768
QY 414 SerAspGlyAlaGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
Db 769 -----GACTATCGCAGCATCTG----- 789
QY 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer 453
Db 790 -----CGTGAATCGGCAT----- 804

```

454	MetArg	GlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGly	473
805	----	 -ACCCCTG-	810
474	GlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGlu	493	
811	----	: : : : : -GAAAGGATAACAATAA--	828
494	GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer	513	
828	----		828
514	IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla	533	
828	----		828
534	AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp	553	
829	----	: : : : : -GTCTCCTACGGCGCTGTGGCT-	852
554	AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe	573	
852	----		852
574	GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPhe	593	
853	----	: : : : : ---CCCCAGGTGAACCGATTGTGATCATGTGCATGTGCTGCCCATATTCACCACTGG	909
594	AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla	613	
910	TGTGGATTACGTCATCTTTACTACGCGCAGGAGATTTTGCTCGCGGGG----	963	
614	ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu	633	
964	-----	: : : : : ---TTTCGATATTAAACACACACGCTGAATCGATCGTGCAGCACGGCGTC	1008
634	ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly	653	
1009	GTTAATCTGGTCTTACCATTGCG-----	1059	
654	ArgArg----PheLeuLeuLeuGlyThrIleProIleLeuLeuAlaSerLeuValLeu	672	
1060	CGCCGTAAATAATGCTGTGCTTGGCGCTCGGGATTG-----	1110	
673	ValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSer	692	
1111	CTGATCGCGCGCTTAGCCATAGGCGCATATGAGGTGCGCGGTACTG-----	1164	
693	ValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys	712	
1165	GTGTGCGCGCGATTGCTATTATGCTGTGACCCCTGCGCGCGGTACCTCGGTCTGCT	1224	
713	AlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPhe	732	
1225	CGCAGATTTTTCCCAACCCGCTACGTGACNCTCGCGATGCTTTAGGTACCTGGCGCTG	1284	
733	TripleGlyAspIleLeuValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeu	752	
1285	TGGATCGCGCTGTTTCCCTGTTAAACCTATACCTTCCCGCTGCTTAATCGCGGTCTCGGCGCG	1344	
753	AlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeu	772	
1345	GCAGGAGCTTCTCGTGHATGTHATTATTTGGCTGCGCGGTACCTCTATATCTCGCGC	1404	
773	LysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAla	790	
1405	AACGTCGCGGAAAACGAAAGGTATATCACCCCTCGAAGCGCTGGAAGACGCTGGCG	1458	

RESULT 15
US-09-489-039A-2378
; Sequence 2378, Application US/09489039A
; Patent No. 6510836
; GENERAL INFORMATION:

```

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2378
; LENGTH: 1566
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2378

Alignment Scores:
Pred. No.: 8.92e-35
Score: 426.00
Length: 1566
Percent Similarity: 34.24%
Matches: 163
Best Local Similarity: 22.24%
Conservative: 88
Query Match: 10.45%
Mismatches: 176
Indels: 306
DB: 4
Gaps: 20

```

US-10-051-909-32 (1-800) x US-09-489-039A-2378 (1-1566)

[illegible][illegible]

```

QY 299 AlaAlaAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGlu 318
Db 793 CGG-----CAGAGGAG-----804
QY 319 GluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeu 338
Db 804 -----804
QY 339 SerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIle 358
Db 805 -----TACGATGACATC 816
QY 359 ValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSer 378
Db 817 TCTACCTGATC-----828
QY 379 ThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGlu 398
Db 829 -----AAAAATTGAA 837
QY 399 GlnTrpAspGluGluAsnLeuHisArgAspGluGluTyrAlaSerAspGlyAlaGly 418
Db 838 -----GCTGAT 843
QY 419 GlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGlu 438
Db 844 AATAAGTAC-----AGCGCCAGAGACCTTTGCG---873
QY 439 GlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgGlnThr 458
Db 873 -----873
QY 459 LeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyTrpGlnLeu 478
Db 873 -----873
QY 479 AlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArg 498
Db 873 -----873
QY 499 ValTyrLeuHisGlnGluGlyValProGlySerArgGlySerIleValSerLeuPro 518
Db 874 -----ACGATTGTCAAAACACCC 891
QY 519 GlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGln 538
Db 891 -----891
QY 539 SerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHis 558
Db 891 -----891
QY 559 ProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArg 578
Db 892 -----TGG-----ATCCTC 900
QY 579 ArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGly 598
Db 901 AAAATTTTCTGGTGGCATTACCTGGCGCCGACCTGCAGCAACACCGCGTTAACGTC 960
QY 599 ValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLys 618
Db 961 ATTATGTATTACGGTACGGAAATCCTGAGTCCGCGCGTTTC-----1002
QY 619 PheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMet 638
Db 1003 -----TCAGCGGTACTTCGCTGATCTGCAACGTCGACGCGTGTCTTCTCC 1050
QY 639 Leu-----ProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArg---655
Db 1051 GTGGCGGTATGCTGATCGGG---GTACTGTCTCTTGTGCGACCGCTTCAAACGTAAGACA 1107
QY 656 -----PheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuVal 670

```

```

Db 1108 ATCATCATCTACGGTTTCGCCATTATGCAACCTCGCATCTGATATATCGCGCTGTCGAT 1167
QY 671 IleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThr 690
Db 1168 TACACCTCGTGGGT-----GACCTGAAAGCTACCGCATCTGGCTGCTGGCGCG 1218
QY 691 ValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIle 710
Db 1219 CTGTTTCGTCGGCGTGATGACGGGTCC-----ATGGGCTTT-----ATTACCTGGGTG 1266
QY 711 LeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPhe 730
Db 1267 GTATTAGCGGAACGTTCCTCCGCTGAATTCGCGGACHTGTCGATGGGATCTCGGTATTT 1326
QY 731 ThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIle 750
Db 1327 TTCATGTGATCATGACGCGTGTGCTCAGCTATCTGTTCCGCTACTGCAGGCAAACTG 1386
QY 751 GlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheVal 770
Db 1387 GGCCTTGGCGCAGCTCTTTTATCTTCGCGCCATTAACTATTATTAGCCATTCTTTTCGTG 1446
QY 771 PheLeuLysValProGluThrLysGlyMetProLeuGlu 783
Db 1447 GTCTTTGGCTTGGCGGAGACCTCCCAATAAATCGCTTCAG 1485

```

Search completed: October 13, 2004, 23:30:06
 Job time : 239.106 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 13, 2004, 16:55:37 : Search time 1141.17 Seconds
(without alignments)
3566.233 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGLVAVQTFTFDLDR.....PLEVITEFVAGAKQAQAAKA 800

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3372238 seqs, 2543544081 residues

Total number of hits satisfying chosen parameters: 6744476

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DBV=xlh
-Q=/cgn2_1/USFTO.spool.h/US10051909/runat_13102004_123339_20010/app.query.fasta_1.1678
-DB=Published Applications NA -QFMT=fastp -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS-human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-MAXLEN=2000000000 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DGPLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
19: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
20: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result

%

Query

No.	Score	Match	Length	DB	ID	Description
1	4075	100.0	2777	13	US-10-051-909-31	Sequence 31, Appl
2	4059	99.6	2908	15	US-10-310-154-350	Sequence 350, App
3	3517	86.3	2824	13	US-10-051-902-1	Sequence 1, Appli
4	3517	86.3	2824	13	US-10-051-902-1	Sequence 1, Appli
5	2869	70.4	2031	16	US-10-425-114-3261	Sequence 3261, A
6	2835.5	69.6	2769	17	US-10-437-963-93472	Sequence 93472, A
7	2674	65.6	2601	13	US-10-051-902-7	Sequence 7, Appli
8	2674	65.6	2601	13	US-10-051-909-7	Sequence 7, Appli
9	2584	63.4	2190	9	US-09-938-842A-1315	Sequence 1315, Ap
10	2584	63.4	2190	11	US-09-938-842A-1315	Sequence 1315, Ap
11	2571	63.1	2671	16	US-10-424-599-10324	Sequence 10324, A
12	2543	62.4	2405	16	US-10-425-114-7958	Sequence 7958, Ap
13	2483.5	60.9	2573	16	US-10-424-599-108435	Sequence 108435, A
14	2442.5	59.9	2271	16	US-10-425-114-29146	Sequence 29146, A
15	2275	55.8	2205	9	US-09-938-842A-2254	Sequence 2254, Ap
16	2275	55.8	2205	11	US-09-938-842A-2254	Sequence 2254, Ap
17	2245.5	55.1	2432	17	US-10-437-963-47424	Sequence 47424, A
18	1963	48.2	1959	17	US-10-437-963-52959	Sequence 52959, A
19	1763	43.3	2231	17	US-10-437-963-70909	Sequence 70909, A
20	1670.5	41.0	1692	13	US-10-051-902-9	Sequence 9, Appli
21	1670.5	41.0	1692	13	US-10-051-909-9	Sequence 9, Appli
22	1545	37.9	3205	16	US-10-424-599-77967	Sequence 77967, A
23	1424	34.9	1487	13	US-10-051-902-13	Sequence 13, Appl
24	1424	34.9	1487	13	US-10-051-909-13	Sequence 13, Appl
25	1392	34.2	1313	17	US-10-767-701-12595	Sequence 12595, A
26	1252.5	30.7	1289	16	US-10-424-599-42000	Sequence 42000, A
27	1035	25.4	1412	16	US-10-260-238-747	Sequence 747, App
28	943	23.3	1259	16	US-10-424-599-101136	Sequence 101136, A
29	943	23.1	1009	13	US-10-051-902-15	Sequence 15, Appl
30	943	23.1	1009	13	US-10-051-909-15	Sequence 15, Appl
31	813	20.0	761	16	US-10-425-114-35910	Sequence 35910, A
32	793	19.5	650	16	US-10-425-114-3864	Sequence 3864, Ap
33	663	16.3	435	17	US-10-767-701-18822	Sequence 18822, A
34	620	15.2	711	17	US-10-767-701-12614	Sequence 12614, A
35	620	15.2	751	16	US-10-260-238-2683	Sequence 2683, Ap
36	615.5	15.1	2253	17	US-10-437-963-49329	Sequence 49329, A
37	604	14.8	674	17	US-10-437-963-60088	Sequence 60088, A
38	600	14.7	778	16	US-10-260-238-2684	Sequence 2684, Ap
39	587.5	14.4	1374	15	US-10-369-493-47011	Sequence 47011, A
40	580	14.2	583	17	US-10-021-323-11734	Sequence 11734, A
41	562	13.8	870	13	US-10-051-902-5	Sequence 5, Appli
42	562	13.8	870	13	US-10-051-909-5	Sequence 5, Appli
43	554	13.6	627	16	US-10-425-114-12008	Sequence 12008, A
44	552	13.5	659	16	US-10-260-238-2685	Sequence 2685, Ap
45	549	13.5	1898	16	US-10-424-599-96711	Sequence 96711, A

ALIGNMENTS

RESULT 1

US-10-051-909-31
; Sequence 31, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: B01163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 31
; LENGTH: 2777
; TYPE: DNA
; ORGANISM: Zea mays

Db 1992 ACAATTCCTGATAGCATCTAGTTATCTCTGTTGTTGTCCTCAATCTAATGATTG 2051
 Qy 681 GlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPhe 700
 Db 2052 GGTACACTAGCCATGCTTCTCTCCACCGTCAGTGTATGCTACTCTCTGCTTC 2111
 Qy 701 ValMetGlyPheGlyProIleProAenIleLeuCysAlaGluIlePheProThrArgVal 720
 Db 2112 GTTATGGGATTGGTCCCATCCCAATTTATGTGCAGAGATCTTTCCAAACCAAGGTT 2171
 Qy 721 ArgGlyLeuCysIleAlaIleCysAlaPheThrPheThrIleGlyAspIleIleValThr 740
 Db 2172 CGTGGGCTCTGTATGCCATTTGTGCTTTACATCTGGATCGGAGATATCATCGTACC 2231
 Qy 741 TyrSerLeuProValMetLeuAenAlaIleGlyLeuAlaGlyValPheSerIleTyrAla 760
 Db 2232 TACAGCTTCTCTGTGTGCTGATGCTATGTGACTGGCGGGTGTTCAGCATATATGCA 2291
 Qy 761 ValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMet 780
 Db 2292 GTCGTATGCTTGATTTCCTTTGTGTTCGCTTCTTAAAGTCCCTGAGACAAAGGGATG 2351
 Qy 781 ProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLysAla 800
 Db 2352 CCCCTTGAGGTATATACCGNATTTCTTGAGTTGGTGGCGAAGCAAGCGGCTGCAAAAGCC 2411

RESULT 2

US-10-310-154-350
 ; Sequence 350, Application US/10310154
 ; Publication No. US20030233670A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Edgerton, Michael D
 ; APPLICANT: Chometon, Paul S.
 ; APPLICANT: Adams, Thomas H
 ; APPLICANT: Ruff, Thomas G.
 ; APPLICANT: Agarwal, Ameeta K.
 ; APPLICANT: Ahrens, Jeffrey E.
 ; APPLICANT: Ball, James A.
 ; APPLICANT: Banu, G.
 ; APPLICANT: Bell, Erin
 ; APPLICANT: Boddupalli, Raghava
 ; APPLICANT: Deikman, Jill
 ; APPLICANT: Deng, Molian
 ; APPLICANT: Dong, Jinzhao
 ; APPLICANT: Duff, Stephen M.
 ; APPLICANT: Galligan, Meghan M.
 ; APPLICANT: Hinchey, Brenda S.
 ; APPLICANT: Huang, Shishieh
 ; APPLICANT: Johnson, G. Richard
 ; APPLICANT: Jung, Vincent
 ; APPLICANT: Kretzmer, Keith A
 ; APPLICANT: Laccetti, Lucille B.
 ; APPLICANT: Lai, Chao-Qiang
 ; APPLICANT: Lee, Gary
 ; APPLICANT: Lin, Jie-Yi
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Lu, Bin
 ; APPLICANT: Luethy, Michael M.
 ; APPLICANT: Lund, Adrian
 ; APPLICANT: Madison, Linda L.
 ; APPLICANT: Malloy, Kathleen A.
 ; APPLICANT: McKiel, Christine L.
 ; APPLICANT: Miller, Philip W.
 ; APPLICANT: Padmavathi, Manchikanti
 ; APPLICANT: Parnell, Laurence D.
 ; APPLICANT: Start, William G.
 ; APPLICANT: Tennesen, Dan
 ; APPLICANT: Vidya, K.R.
 ; APPLICANT: Wang, Haiyun
 ; APPLICANT: Xin, Zhaunguo
 ; APPLICANT: Xu, Nanfei
 ; APPLICANT: Yang, Chunzhi
 ; APPLICANT: Zeng, Xiaoping

; APPLICANT: Zhang, Qiang
 ; APPLICANT: Zhao, Yajuan
 ; APPLICANT: Zhou, Li
 ; TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
 ; FILE REFERENCE: 38-15(52796)B
 ; CURRENT APPLICATION NUMBER: US/10/310,154
 ; CURRENT FILING DATE: 2002-12-04
 ; PRIOR APPLICATION NUMBER: 60/337,358
 ; PRIOR FILING DATE: 2001-12-04
 ; NUMBER OF SEQ ID NOS: 736
 ; SEQ ID NO 350
 ; LENGTH: 2908
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (331)..(2565)
 ; OTHER INFORMATION:
 US-10-310-154-350
 Alignment Scores:
 Pred. No.: 0 Length: 2908
 Score: 4059.00 Matches: 797
 Percent Similarity: 99.75% Conservative: 1
 Best Local Similarity: 99.62% Mismatches: 2
 Query Match: 99.61% Indels: 0
 DB: 15 Gaps: 0
 US-10-051-909-32 (1-800) x US-10-310-154-350 (1-2908)
 Qy 1 IleArgSerGlySerThrLeuAlaValGlnThrProPheThrProAspLeuAspArg 20
 Db 166 ATTCGGAGCGGCTTTGGCTTCAGTCCAGACGCCCTTACCCCTGATCTGACCGGAGG 225
 Qy 21 GluArgLeuLeuProSerValValLeuAlaLeuProGlyProLeuProAlaSerCys 40
 Db 226 GAGCGGCTCCTTCGCTCAGTTGTTCTTGCTTTGGCTGGCGCTCTTCGCCCTGCTTCGTGT 285
 Qy 41 SerSerGlnGluProValThrSerAspAspIleLeuGluAspLysMetSerGlyAlaVal 60
 Db 286 TCTTCACAGAGCGCGTGACCTCGACGATATCTTGGAGGACAGATGTTCGGGGGCTGTT 345
 Qy 61 LeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyrAspAsnAlaThrIle 80
 Db 346 CTTGTGCGCATAGTCGCTCCATCGGCAATTAITTCAGGCGGTGGACAATGCCACCATC 405
 Qy 81 AlaAlaAlaValLeuTyrIleIleIleIleIleIleIleIleIleIleIleIleIleIle 100
 Db 406 GCAGCTGCTGTCTGTATATATAAGAGGAATTCATTTGCAAAATGACCCACTGTGGAG 465
 Qy 101 GlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrPheSerGlyPro 120
 Db 466 GGACTAATTTGTCAATGTCATCTATCGCGCCACCATCGTTACTACATTTCCGGGCA 525
 Qy 121 LeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePhe 140
 Db 526 TTATCAGACTGCTGATGGCGCGCCCTATGCTTATTTCTCTTCAATTTCTACTTCTTC 585
 Qy 141 SerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeuAlaArgPheVal 160
 Db 586 ACGGCGCTCATCATGCTATGTTCTCTTAATGCTATGCTGCTGCTGCTGCTGCTGCTGCT 645
 Qy 161 AspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAla 180
 Db 646 GATGGATTTGGTATTGGCTTGGCTGTCAAGCTTGGCTTGTGTATACATTTCAAAATAGCC 705
 Qy 181 ProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMet 200
 Db 706 CCTTCGGAGATTAGAGGTTTCTGTAATACATCTACCAATTCAGTGGATTCAGAGGAATG 765
 Qy 201 PheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSerProAspThrArgIle 220
 Db 766 TTCTTGTCTACTGTCATGTTGTTGGATGTCCTTGTGCGCATCACCGGATGGAGAATT 825

QY 221 MetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeu 240
DB 826 ATGCTTGCTGCTCGGATACCTTCATTGTTCTTCTTGTTGACAAATTTTATCTT 885
QY 241 ProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysValLeu 260
DB 886 CCTGAATCTCCAAAGATGGCTGCTTTAGCAAGGTCGGATCGCAGAGCAAAAAGGTGTTG 945
QY 261 GlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeuLeuGluGlyLeu 280
DB 946 CAAAGTTACGGGGGAAAGAGATGCTCAGGTGAATTTGCTCTTCTTCGAGGGTTG 1005
QY 281 GluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAla 300
DB 1006 GAGGTTGGAGGACACATTCATTGAAGAGTACATCAATTTGGACCTGCCACCGAGGAGCC 1065
QY 301 AspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGly 320
DB 1066 GATGATCTTGTTACTGACGGTGATAGGAACAATCACACTTATGGGCTGAGAGGCT 1125
QY 321 GlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeu 340
DB 1126 CAGTCATGGATTCTCGACCTTCAAGGAGCCAGCATGCTTGGAGTGTGCTTCTCTT 1185
QY 341 AlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThr 360
DB 1186 GCATCTCGTCATGGAGCATGGTGAACAGAGTGTACCCCTTATGGATCCGATTGTGACA 1245
QY 361 LeuPheGlySerValHisGluAsnMetProGlnAlaGlyLysSerMetArgSerThrLeu 380
DB 1246 CTTTGTGGTAGTCTCCATGAGATATGCCTCAAGCTGGAGGAATATGAGGACACATTTG 1305
QY 381 PheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluInTrp 400
DB 1306 TTTTCAAACCTTTGGAAAGTATGTTCACTGTGCACAGATCAGCATGCCAAAATGAGCAGTGG 1365
QY 401 AspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyLys 420
DB 1366 GATGAGAGATCTTATAGGATGACGAGGATGACGATCTGATGGTGCAGAGGTGAC 1425
QY 421 TyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLys 440
DB 1426 TATGAGGACAACTCTCCATAGCCCATTTGCTGTCCAGGACGCAACAGTCCGGAAGGAAG 1485
QY 441 AspIleValHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu 460
DB 1486 GACATTTGTGCACCATGTCACCGTGAAGTGTCTTTCAGCATGAGAGGCAAGCCTCTTTA 1545
QY 461 GlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyTrpGlnLeuAlaTrp 480
DB 1546 GGGGAGGTTGGAGATGTTGAGCAGCACTGATATCGTGGGGATGGCAGCTTGGCTGG 1605
QY 481 LysTrpSerGlnLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyr 500
DB 1606 AAATGTGTAGAGGAAGAGTGTAGATGTGTAGAAAGAGGAGGTTCAAAAGAGTCTAC 1665
QY 501 LeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly 520
DB 1666 TTGCACCAAGAGGAGTCTCTGCTCAAGAGGGGCTCAATTTTCACTTCCCGGTGGT 1725
QY 521 GlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAla 540
DB 1726 GCGCATGTTCTTGAGGGTATGAGTTGTACATGCTGCTGCTTTTATAGTACGTCAGCACA 1785
QY 541 LeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSer 560
DB 1786 CTTTCTCAAAGGCTCTTCTGTCACCAAGCATGTCAGATGCTGCTGCTTCACTTCACTCT 1845
QY 561 GluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgAla 580
DB 1846 GAGGTAGTGCCTCAAGAGGTTCACTGTTGAAAGATTTGTTGAAACCTGAGTGGGGTGGCC 1905

QY 581 LeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeu 600
DB 1906 CTGTTAGTCGGTGTGGAAATTCAGATCCTTCAACAGTTTCTGGAATAAAACGGTGTCTG 1965
QY 601 TyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGly 620
DB 1966 TACTATATCCCAAAATCTTGGCAAGCTGGTGGCAGTTATCTTCCAAATTTGGT 2025
QY 621 LeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuPro 640
DB 2026 CTCAGCTCGGCATCAGCATCCTTTCATCAGTTCTCTCACTACCTTACTAAATGCTTCT 2085
QY 641 CysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGly 660
DB 2086 TGCATTGGCTTTCATGCTGCTTATGATCTTTCGGAAAGAGGTTTTGCTGCTAGGC 2145
QY 661 ThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeu 680
DB 2146 ACATTTCCATCTTGTATGATGATCTCTAGTTATCTCTGGTGTGTCATCTAATTTGTTG 2205
QY 681 GlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPhe 700
DB 2206 GGTACACTAGCCCATGCTTTCCTCCACCGTCAGTGTATCTGCTACTTCTCTGCTTC 2265
QY 701 ValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgVal 720
DB 2266 GTTATGAGATTTGGTCCATCCCAATTTTATGTGACAGATCTTCCACAGGGTT 2325
QY 721 ArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThr 740
DB 2326 CGTGGCTCTGATTTGCAATTTGCTTACATTTCTGATCGGAGATATCATCTGTCACC 2385
QY 741 TyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAla 760
DB 2386 TACAGCTTCTGTGATGTGTAATGCTATTTGGACTGCGGGTGTTCAGCATATATGCA 2445
QY 761 ValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMet 780
DB 2446 GTCGTATGCTTGAATTTCTTTGTTGCTCTTCTTAAAGTCTCTGAGACAAAGGGATG 2505
QY 781 ProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLysAla 800
DB 2506 CCCCTTGAAGTTATACCGAATTTTTCAGTTTGGTGGGCAAGCGGCTGCAAAAGCC 2565
RESULT 3
US-10-051-902-1
; Sequence 1, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; PRIORITY FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 2824
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (29)
; NAME/KEY: unsure
; LOCATION: (622)
; NAME/KEY: unsure
; LOCATION: (636)
; NAME/KEY: unsure


```

; LOCATION: (638)
; NAME/KEY: unsure
; LOCATION: (669)
; NAME/KEY: unsure
; LOCATION: (771)
; NAME/KEY: unsure
; LOCATION: (822)
; NAME/KEY: unsure
; LOCATION: (856)
; NAME/KEY: unsure
; LOCATION: (889)
; NAME/KEY: unsure
; LOCATION: (896)
; NAME/KEY: unsure
; LOCATION: (944)
; NAME/KEY: unsure
US-10-051-902-1

Alignment Scores:
Pred. No.: 0 Length: 2824
Score: 3517.00 Matches: 684
Percent Similarity: 95.18% Conservative: 27
Best Local Similarity: 91.57% Mismatches: 34
Query Match: 86.31% Indels: 2
DB: 13 Gaps: 2

US-10-051-909-32 (1-800) x US-10-051-902-1 (1-2824)
QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuGlnGlyTyr 75
Db 238 ATGGGGGGCCCGTGTATGGTGGCCATCGCGGCCCTCTATCGGAACCTGCTGAGGGCTGG 297
QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysGluPheGlnLeuGlnAsn 95
Db 298 GACAAATCGACAATTGCTGAGCGCTCTGTACATAAAGAGGAATTCAACCTGCAGAGC 357
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db 358 GAGCCCTCTGATCGAAGGCCCTCATCGTGGCCATGTTCTCATTTGGGGCAACATCATACA 417
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgProMetLeuIleLeuSerSer 135
Db 418 ACATCTCCGGGGCAAGGGCTGACTCGGTGGTAGAGGCCCATGCTGGTGGCTCGCT 477
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu 155
Db 478 GTCTCTACTTCTGATGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 537
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
Db 538 CTCGCAAGGCTCATGATGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 597
QY 176 IleSerGluIleAlaProSerGluIleArg--GlyLeuLeuAsnThrLeuProGlnPhe 194
Db 598 ATCTCCGAAACTGCACCGCACANATTCTTGGGGCTGNTNGAACAACGTTGGCGAGTTC 657
QY 195 SerGly---SerGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213
Db 658 ATGGGGTTCAGNGAGGAGATGTTCTCTCTACTGATGGTGGTGGTGGTGGTGGTGGTGGT 717
QY 214 ProSerProAspTyrPheMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233
Db 718 CCCAAACCTGATGGAGGCTATGCTTGGAGTTCCTCGATCCGTCACCTTATNACTTT 777
QY 234 GlyLeuThrIlePheTyrLeuProGluSerProArgTyrPheValSerLysGlyArgMet 253
Db 778 GGACTGACTCTCTTCTACTTGGCTGATCACCACCAAGGTGGCTGTNNAGCAAGAGGATG 837
QY 254 AlaGluAlaLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeu 273
Db 838 GCGAGCGGAGAGATGNTGCAAGGCTCGGGGAGAGAGATGTCATCANGGAGANG 897
QY 274 SerLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluTyrIleIle 293
Db 897

```

```

Db 898 GCTCTTCTAGTTGAAGGTTTGGGGTGGTAAAGATACACGTAATTTNAGAGTACATCAT 957
QY 294 GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThr 313
Db 958 GGACTGCTCCACCGAGGAGCGATGATCTTTAACTACCGTGTATAGGAACAATCACA 1017
QY 314 LeuTyrGlyProGluGluGlyGlnSerTyrIleAlaArgProSerLysGlyProIleMet 333
Db 1018 CTTTATGGGCTGAAAGAGGCCATGATGATTTCTGATCTTCAAGGATGATGATGATGAT 1077
QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
Db 1078 CTTGGAAGTGTGCTTCTTCTGATCTCGTCATGGAGCATGGTGAACAGAGGTGATCCC 1137
QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
Db 1138 CTTATGGATCGATGTGACACTTTTGGTAGTGTCCATGAGAATATGCTTCAAGCTGGA 1197
QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db 1198 GGAAGTATGAGGAGCACATTTGTTCCAAACTTTGGAAGTATGTTTCAGTGTCCAGATCAG 1257
QY 394 HisAlaLysAsnGluGlnTyrAspGluGluAsnLeuHisArgAspAspGluGlyTyrAla 413
Db 1258 CATGCCAAAATGAGCATGGGATGAGAGAAATCTTATAGGGATGACGAGGATGACGA 1317
QY 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
Db 1318 TCTGATGTGAGAGGATGACTATGAGGACATCTCCATAGCCCATTCCTGCTCCAGGCGAG 1377
QY 434 AlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGlySerAlaLeuSer 453
Db 1378 GCAACAGGTGGCGAGGAGGAGGACATGTCACCATGTCACCGTGGAAAGTCTTTGAGC 1437
QY 454 MetArgArgGlnThrLeuLeuGlyGluGlyValAspGlyValSerSerThrAspIleGly 473
Db 1438 ATGAGAAGCAAAACCTTTAGGGAGGAGTGGAGATGGTGGAGCAGCATGATATCGT 1497
QY 474 GlyGlyTyrGlnLeuAlaTyrPheSerGluLysGluGlyGluAsnGlyArgLysGlu 493
Db 1498 GGGGATGGCAGCTTGTCTGGAATGTCAGAGAAGGAGGTGAGATGTTGTAAGAAGAA 1557
QY 494 GlyGlyPheLysArgValTyrLeuHisGlnGlyValProGlySerArgArgGlySer 513
Db 1558 GGTGGTTCCTTCCCGTGGTGGGAGTCTTCCACCAAGAGGAGTCTCTGGCTCAAGAAGGGTCA 1617
QY 514 IleValSerLeuProGlyGlyValPheGlySerGluPheValHisAlaAla 533
Db 1618 ATTGTTTCACTTCCCGTGGTGGGAGTCTTCTGAGGGTAGTGAGTTTGTACATGCTGCT 1677
QY 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553
Db 1678 GCTTTAGTAACTAGTCAGCAGCACTTTCTCAAAAGGCTCTTGTGAACCCAGCATGTCAGAT 1737
QY 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTyrLysAspLeuPhe 573
Db 1738 GCTGCCATGGTTCACCCATCTGAGGTAGTCCCAAGGTTCCAGTTGGAAAGATTGTTT 1797
QY 574 GluProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPhe 593
Db 1798 GAACCTGGAGTGAGGCGCTGTTAGTGGTGTGGATTCAGATTCATGATTCAGATGTTT 1857
QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGlnAlaGlyValAla 613
Db 1858 GCTGGAATAAAGCGTGTCTGTACTATACCCCAAAATCTTTGAGCAGGCTGGGTGGCA 1917
QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
Db 1918 GTTATTTCTTCCAAATTTGGTCTCAGCTCGGCATCAGCATCCATCTTGATCAGTTCTCTC 1977
QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
Db 1978 ACTACCTTACTAAATGCTTCTTGTGCTTGGCTTTCCTGCTTGGCTTTCCTGATCTTCCGGA 2037

```

```

Qy 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuLeuAlaSerIleValIleLeuVal 673
Db 2038 AGAAGGTTTTTGTGTAGGACAAATCCAAATCTTGATAGCATCTCTAGTTATCTCTGGTT 2097
Qy 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
Db 2098 GTGTCCAAATCAATTAATTTGGGTACACTAGCCATGCTTTTCTCCACCATCAGTGT 2157
Qy 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIlePheProAsnIleLeuCysAla 713
Db 2158 ATCGTCTACTCTCTGCTGCTCGTTATGGGATTTGGTCCCATCCCAACATTTTATGTGCA 2217
Qy 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
Db 2218 GAGATCTTTCCAAACAGGGTCTGGCCCTCTGATTCGCAATTTGGCCCTTACATCTGG 2277
Qy 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
Db 2278 ATCGGAGATATCATCTGACCTACAGCTTCCTGTGATGCTGAATGCTATTGGACTGGCG 2337
Qy 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773
Db 2338 GGTGTCTTTCAGCATATATGACGTCGTATGCTGATTTCTTGTGTTCTCTCTCTTAAG 2397
Qy 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
Db 2398 GTCCCTGAGACAAAGGGGATGCCCTTGAGTTATTACCGAATTTCTTTCAGTTGGTGGC 2457
Qy 794 LysGlnAlaAlaLysAla 800
Db 2458 AAGCAAGCGGCTGCAAAAGCC 2478

RESULT 4
US-10-051-909-1
; Sequence 1, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 2824
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (29)
; NAME/KEY: unsure
; LOCATION: (622)
; NAME/KEY: unsure
; LOCATION: (636)
; NAME/KEY: unsure
; LOCATION: (638)
; NAME/KEY: unsure
; LOCATION: (669)
; NAME/KEY: unsure
; LOCATION: (771)
; NAME/KEY: unsure
; LOCATION: (822)
; NAME/KEY: unsure
; LOCATION: (856)
; NAME/KEY: unsure

```

```

; LOCATION: (889)
; NAME/KEY: unsure
; LOCATION: (896)
; NAME/KEY: unsure
; LOCATION: (944)
US-10-051-909-1
Alignment Scores:
Pred. No.: 0 Length: 2824
Score: 3517.00 Matches: 584
Percent Similarity: 95.18% Conservative: 27
Best Local Similarity: 91.57% Mismatches: 34
Query Match: 86.31% Indels: 2
DB: 13 Gaps: 2
US-10-051-909-32 (1-800) x US-10-051-909-1 (1-2824)
Qy 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
Db 238 ATGGGGGGCGCGGTGATGCGCCATCGCGCCCTCTATCGGCAACTTGCTGCGAGGGCTGG 297
Qy 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
Db 298 GACAAATCGCAATTTGCTGGAGCGCTCTGTATATAAAGAGAGAAATTCACCTGCGAGC 357
Qy 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db 358 GAGCCTCTGATCGAAGGCCCTCATGTCGCCATGTTCTCATTTGGGGCAACAGTCATCACA 417
Qy 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
Db 418 ACATCTCCGGGGCCAGGGCTGACTGCGTTGGTAGAGGCCCATGTGTGGCTCGGCT 477
Qy 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu 155
Db 478 GTCTCTTACTTGTGTCAGTGGCTGGTGATGCTTTGGGGCCAAATTTGTGTACATCTTGCTC 537
Qy 156 LeuAlaArgPheValAspGlyPheGlyIleGlyValAlaValThrLeuValProLeuTyr 175
Db 538 CTCGCAAGGCTCATTTGATGGTTGCGGTATCGGTTGGCGGTACACCTTGTTCCTCTCTAC 597
Qy 176 IleSerGluIleAlaProSerGluIleArg--GlyLeuLeuAsnThrLeuProGlnPhe 194
Db 598 ATCTCCGAAACTGCACGCGACAGANATTTCTGGGGCTGNTNGAACACGTTGCCGAGTTC 657
Qy 195 SerGly---SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213
Db 658 ATTGGGGTCAGNGGAGGAGATGTTCTCTCTACTGTCATGTTGGTGTGGGATGTCCTCATG 717
Qy 214 ProSerProAspTrpArgIleMetLeuGlyValIleAlaIleProSerLeuPhePhePhe 233
Db 718 CCCAACTGATTTGGAGGCTCATGCTGGAGTTCTGTGATCCCGTCACTTATTAATTT 777
Qy 234 GlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMet 253
Db 778 GGACTGACTGCTTCTTACTTGGCTGAATCACAAGGTGGTCTGTNAGCAAGGAAGGATG 837
Qy 254 AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeu 273
Db 838 GCGGAGCGGAGAGAGATGTCGAAAGGCTCGGGGAGAGAGAGATGTCCTCANGGGAGANG 897
Qy 274 SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIle 293
Db 898 GCTCTTCTAGTTGAAGGTTTGGGGGTCCGTAAGATACACGATTTTNNAGAGTACATCA 957
Qy 294 GlyProAlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGlnIleThr 313
Db 958 GGACCTCCACCGGAGGCGCGATGATCTTGTAACTACGCGTATAGGAACAAATCACA 1017
Qy 314 LeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMet 333
Db 1018 CTTTATGGGCTGAGAAAGGCCAGTCATGATGATGCTCGACCTTCTTAAGGGACCCATCATG 1077

```

QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
 DB 1078 CTTGGAAAGTGCTTTCTCTTCATCTCTCATGGAGCATGGTAAACACAGAGTGATACC 1137
 QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
 DB 1138 CTTATGGATCCGATTTGACACTTTTGGTAGTGTCATGAGAAATATGCTCAAGCTGGA 1197
 QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
 DB 1198 GGAAGTATGAGGAGCACTTTTCCAACTTTGGAAAGTATGTTCAAGTGTACAGATCAG 1257
 QY 394 HisAlaIysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTrpAla 413
 DB 1258 CATGCCAAAAATGAGCAGTGGGATGAAGAGAAATCTTCATAGGATGACGAGGATACGCA 1317
 QY 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
 DB 1318 TCTGATGGTGCAGGAGGTGACTATGAGGACAATCTCCATAGCCCAATGCTGTCCAGGCGAG 1377
 QY 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer 453
 DB 1378 GCAACAGGTGCGAAGGAAGGACATTTGTGCACCATGTCACCGTGGAAAGTCTTTGAGC 1437
 QY 454 MetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGly 473
 DB 1438 ATCAGAAGGCAGAACCTCTTAGGGAGGTGGAGATGCTGTGAGCAGCACTGATATCGGT 1497
 QY 474 GlyGlyTyrGlnLeuAlaTrpIleValHisGlyGluGlyGluGlyGluGlyGluGly 493
 DB 1498 GGGGGATGCGAGCTTCTGGAAATGTCAGAGAGGAAGGTGAGATGTTAGAAAGGNA 1557
 QY 494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer 513
 DB 1558 GGTGGTTTCAAAAGAGTCTACTTGCCACCAAGAGGGAGTTCCTGGCTCAAGAGGGGCTCA 1617
 QY 514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533
 DB 1618 ATTGTTTCTACTTCCCGGTGGCGGATGTTCTTGGGGTAGTGAGTTTGATCATGCTGCT 1677
 QY 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553
 DB 1678 GCTTTAGTAGTCACTCAGTCAGCATTTTCTCAAGGGTCTTGCTGAACCAACGATGTCAGAT 1737
 QY 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe 573
 DB 1738 GCTGCCATGGTTCCACCATCTGAGGTAGCTGCCAAAGGTTCACTGTGGAAAGATTGTGTT 1797
 QY 574 GluProGlyValArgArgAlaLeuValGlyValGlyIleGlnIleLeuGlnPhe 593
 DB 1798 GAACCTGGAGTGAGGGTGCCTTGTAGTCGGTGTGGAAATTCAGATCCTTCAACAGTTT 1857
 QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613
 DB 1858 GCTGTGAATAAACGGTGTCTGTACTATACCCACAAATCTTGAGCAAGCTGCTGTGGCA 1917
 QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerLeu 633
 DB 1918 GTTATTCTTTCCAAATTTGCTCTCAGCTGGCATCAGATCCATCTTGATCAGTTCTCTC 1977
 QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
 DB 1978 ACTACCTTACTAATGCTTCTGCAATGGCTTTGGCCATGCTGCTTATGATCTTCCGGA 2037
 QY 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
 DB 2038 AGAAGGTTTTTGTGTAGGACAAATTCCAATCTTGATGATCTCTAGTTATCTCGGTT 2097
 QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
 DB 2098 GTGTCCAAATCAATTGATTTGGGTACACTAGCCCATGCTTTGCTCTCCACCATCAGTGT 2157
 QY 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713

DB 2158 ATCGTCTACTTCTGCTGCTTATGGGATTTGGTCCCATCCCAACATTTTATGTGCA 2217
 QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
 DB 2218 GAGATCTTCCAAACAGGGTTCGTGGCTCTGTATTCGCAATTTGTGCTTTACATTTCTGG 2277
 QY 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
 DB 2278 ATCGAGATATCATCGTCACCTACAGCTTCTCTGTGATGCTGATGCTATTGGACTGGCG 2337
 QY 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuIys 773
 DB 2338 GGTGTTTTTTCAGCATATATGCACTGCTATGCTTGTATTTCTTTGTTGCTGCTTCCCTTAAG 2397
 QY 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
 DB 2398 GTCCCTGAGCAAAAGGGATGCCCTTGAGGTATTATACCGAATTTCTTTGCAGTTGGTGG 2457
 QY 794 LysGlnAlaAlaLysAla 800
 DB 2458 AAGCAAGCGGTGCAAAAGCC 2478
 RESULT 5
 US-10-425-114-32961
 ; Sequence 32961, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaska, Jack E.
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 32961
 ; LENGTH: 2031
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: UC-ZMFLMOL7045C09_FLI
 US-10-425-114-32961
 Alignment Scores:
 Pred. No.: 2,09e-274 Length: 2031
 Score: 2869.00 Matches: 561
 Percent Similarity: 99.65% Conservative: 1
 Best Local Similarity: 99.47% Mismatches: 2
 Query Match: 70.40% Indels: 0
 DB: 16 Gaps: 0
 US-10-051-909-32 (1-800) x US-10-425-114-32961 (1-2031)
 QY 237 IlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAla 256
 DB 2 ATATTTTATCTCTCTGAAATCTCCAAAGATGGCTGTTAGCAAGGTCCGATGGCAGAGGCA 61
 QY 257 LysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeuLeu 276
 DB 62 AAAAAGGTGTTGCCAAAAGTTACGGGGGAAAGACCATGCTCAGGTGAATTTGCTCTTCTT 121
 QY 277 LeuGluGlyLeuGluValGlyAspThrSerIleGluGluTyrIleIleGlyProAla 296
 DB 122 CTCGAAGGTTTGGAGTTGGAGGAGACACTTCCATTGAAGAGTACATCATTTGGACCTGCC 181
 QY 297 ThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGly 316
 DB 182 ACCGAGGCGCCGATGATCTTTTACTGACGCTGATAGGAACAATACACACTTTATGGG 241

QY 317 ProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySer 336
DB 242 CCTGAAGAAGCCAGTCATGGATTGCTCGACCTTCCAAAGGACCAGCATGCTTGGAGT 301
QY 337 ValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAsp 356
DB 302 GTCTCTTCTCTTCATCTCGTCATGGGAGCATGGTGAAACACAGAGTGATCCCTTTATGGAT 361
QY 357 ProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMet 376
DB 362 CCATTTGTGACACTTTTGGTAGTGTCCATGAGATATGCTCAGCTGAGGAGATATG 421
QY 377 ArgSerThrLeuPheProLeuPheGlySerMetPheSerValThrAspGlnHisAlaLys 396
DB 422 AGGAGCACATTTGTTCCAAACTTTGGAAGTATGTTCCAGTGTCCACAGATCAGCATGCCAAA 481
QY 397 AsnGluGlnTrpAspGluGlnLeuHisArgAspGluGluTrpAlaSerAspGly 416
DB 482 AATGACAGTGGATGAAGAGATCTTCATAGGATGACGAGATGACCATCTGATGT 541
QY 417 AlaGlyGlyAspTrpGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGly 436
DB 542 GCAGGAGGTGACTATGAGACAACTTCCATAGCCCATGCTGTCAGGCGAGCAACAGGT 601
QY 437 AlaGluGlyAspIleValHisGlyHisArgGlySerAlaLeuSerMetArgArg 456
DB 602 GCGAAGGAAGGACATTTGACCATGTCACCGTGGAGGTCTTTGAGCATGAGAAG 661
QY 457 GlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyTrp 476
DB 662 CAAGCTCTTGGGAGGAGTGGAGATGTTGTGAGCAGCACTGATATGCTGGGAGTGG 721
QY 477 GlnLeuAlaTrpLysTrpSerGluLysGluGlyAsnGlyArgLysGluGlyPhe 496
DB 722 CAGCTTGTGGAAATGGTCAGAGAGGAGGTGAGATGGTAGAAGGAAGGTGGTTTC 781
QY 497 LysArgValTrpLeuHisGlnGluGlyValProGlySerArgGlySerIleValSer 516
DB 782 AAAAGAGTCTACTTGCACCAAGGAGTCTTGGCTCAAGAGGGGCTCAATGTTTCA 841
QY 517 LeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaLeuVal 536
DB 842 CTTCCCGGTGGTGGCATGTTCTTCAAGGAGTGGATGGTAGATGCTGCTCTTTAGTA 901
QY 537 SerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaLeuMet 556
DB 902 AGTCAGTCAGCACTTTTCTCAAGGGTCTTGTGTAACACCGCATGTCAGATGCTGCCATG 961
QY 557 ValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGly 576
DB 962 GTTCCACCATCTGAGGTAGCTGCAAGGTTCCAGTTGGAAAGATTTGTTTGAACCTGGA 1021
QY 577 ValArgArgAlaLeuValGlyValGlyIleGlnIleLeuGlnPheAlaGlyIle 596
DB 1022 GTAGGCGGTGCTGTGTGTGTGGATTTGAAATTCAGATCCTTCAACAGTTTCTGGAATA 1081
QY 597 AsnGlyValLeuTrpThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeu 616
DB 1082 AACGGTGTCTGTACTATACCCACAAATTTCTTGAAGAGTGTGTGGCAGTTATTTCT 1141
QY 617 SerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrLeu 636
DB 1142 TCCAAATTTGGTCTCAGCTCGCATCAGCATCCATCTTGATCAGTTCCTCACTACCTTA 1201
QY 637 LeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPhe 656
DB 1202 CTAAATGCTTCTTGCATTTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1261
QY 657 LeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValSerAsn 676
DB 1262 TTGCTGCTAGGCACAAATTCATCTTGATAGCATCTCTAGTTATCTGTTGTTGTTGTT 1321

QY 677 LeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyr 696
DB 1322 CTAATTGATTGGGTACATAGCCCATGCTTCTCCACCGCTAGTGTATCGTCTAC 1381
QY 697 PheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePhe 716
DB 1382 TTCTGCTGCTTGGTATGGATTTGGTCCCATCCCAACATTTATGTCAGAGATCTTT 1441
QY 717 ProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheIleGlyAsp 736
DB 1442 CCAACACGGGTGCTGGCCCTCTGATTGCCATTTGTCCTTACATCTCGATCGAGAT 1501
QY 737 IleIleValThrTrpSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPhe 756
DB 1502 ATCATGCTACCTACACGCTTCTGTGATGCTGAATGCTATGACCTGGCGGTGTTTC 1561
QY 757 SerIleTrpAlaValValCysLeuIleSerPheValPheValPheLeuLysValProGlu 776
DB 1562 AGCATATATGACGTCGTATGCTTGAATTCCTTTGTGTCCTTAAAGTCCCTGAG 1621
QY 777 ThrLysGlyMetProLeuGluValIleThrGluPheAlaValGlyAlaLysGlnAla 796
DB 1622 ACAAGGGGATGCCCTTGAGTTATTACCGAATCTTTGCAAGTGTGCGAAGCAAGCG 1681
QY 797 AlaAlaLysAla 800
DB 1682 GCTGCAAAAGCC 1693

RESULT 6
US-10-437-963-93472/c
; Sequence 93472, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 93472
; LENGTH: 2769
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91853C.1
US-10-437-963-93472

Alignment Scores:
Pred. No.: 7,17e-271 Length: 2769
Score: 2835.50 Matches: 562
Percent Similarity: 82.45% Conservative: 77
Best Local Similarity: 72.52% Mismatches: 120
Query Match: 69.58% Indels: 16
DB: 17 Gaps: 7

US-10-051-909-32 (1-800) x US-10-437-963-93472 (1-2769)

QY 30 AlaLeuProGlyProLeuProPro-AlaSerCys-----SerSerG1 43
DB 2684 GCCTTCCAGATTCACGAGCGCTCTCTTCTTAGGGGATCCGAAATCTCGGTGGACG 2625
QY 43 nGluProValThrSerAspAspIleLeuGluAspLysMetSerGlyAlaValLeuVal 63
DB 2624 AGAGACTGTGTGTAAGAT-----TCGCCGCCATGGCGGCGCGCTGCTGTGTCG 2574

QY 63 aIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIleAlaAla 83
Db 2573 CATCGGGCTTCATCGGCACTTCTGTCAGGGCTGGGATATATGCAACATTGCAAGTGC 2514
QY 83 aValLeuTyriIleLysLysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeuI 103
Db 2513 GGTACTGTACATCAAGAGGAATTCAACTTGAGAGCGAGCCCTTATCGAAGGCGCTGAT 2454
QY 103 eValSerMetSerLeuIleGlyAlaThrIleValThrPheSerGlyProLeuSerAs 123
Db 2453 CGTGGCCATGTGCTCATTTGGGGCGACGATCATCAGCGTCTCTGAGAGAGTGGCTGA 2394
QY 123 pSerIleGlyArgProMetLeuIleLeuSerSerIleLeuTyriPhePheSerGlyLe 143
Db 2393 TTTCTTTGGTAGGGCGCCCATGCTGATCGGCTCGGCTGCTCTACTTTTGTAGTGGCT 2334
QY 143 uIleMetLeuTrpSerProAsnValTyriValLeuLeuLeuAlaAaGpHeValAspGlyPh 163
Db 2333 AGTGATGCTTTGGGCCCAAAATGTATGTGTGCTTTGGGAGGCTCAITGACGGGT 2274
QY 163 eGlyIleGlyLeuAlaValThrLeuValProLeuTyriIleSerGluIleAlaProSerG 183
Db 2273 CGGGATCGGTTGGCTGTGTCAGCTGTACCATTTGTACATCTCTGAGACTGCCCGGACGGA 2214
QY 183 uIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeuSe 203
Db 2213 CATCAGAGACTGCTAAACACGCTGCCGAGTTCACTGGGCTCTGGAGGATGTTCTCTTC 2154
QY 203 rTyriCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMetLeuG 223
Db 2153 ATACTGCATGTATTGGCATGTTCCTCATGCCACAGCAGATTGAGGATCATGCTTGG 2094
QY 223 yValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyriLeuProGluSe 243
Db 2093 CGTTCTATCAATACCATCACTTATATACCTTTGCTTGGCATTCCTTTTACTTACCTGAATC 2034
QY 243 rProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLe 263
Db 2033 GCCGAGTGGCTCGTGAGCAAGAGAGATGGCTGAGGCGCAAGCGGTGTGTGCAAGGCT 1974
QY 263 uArgGlyLysAspValSerGlyGluLeuSerLeuLeuLeuGluGlyLeuGluValG 283
Db 1973 GCGTGGNAGAGAGATGTTTCAGAGAAATGGCCCTTCCTGTTGAAGGCTGGGGTTGG 1914
QY 283 yGlyAspThrSerIleGluTyriIleLeuGlyProAlaThrGluAlaAlaAspAsple 303
Db 1913 GAAAGACACAAAATTGAGGATACATAATTGGACCTGATGATGAGCTTGTCTGATGAAGG 1854
QY 303 uValThrAspGlyAspLysGluGlnIleThrLeuTyriGlyProGluGluGlyGlnSerTr 323
Db 1853 GCTGGCTCCA--GATCCAGAGAGATCAAACTGTATGTTCTCGAAGAGGCTTATCGTG 1797
QY 323 pIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerAr 343
Db 1796 GGTGGCCGCTCTGTTCAACGGCAAGTGCACTTGAAGTGCATTAGGTCTCATCTCTCG 1737
QY 343 gHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheG 363
Db 1736 TCATGGTAGTATGCTAGTCAGGGAAGAGCCCTTGTGGATCCCTGTGTACCCCTTTTGG 1677
QY 363 ySerValHisGluAsnMetProGlnAlaGlyLysSerMetArgSerThrLeuPheProAs 383
Db 1676 AAGTGTCCATGAGAAGATGCTGAGATAATGGGAAGCATCGGAGGACATTGTTTCTCTAA 1617
QY 383 nPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGlu 403
Db 1616 CTTTGGCAGATGTTTATGTGTGGCGGAACGACGACAGCTAAAGGTGATTTGGGATGCTGA 1557
QY 403 uAsnLeuHisArgAspAspGluGluTyriAlaSerAspGlyAlaGlyLysAspTyriGluAs 423
Db 1556 GAGT--CAACGGAGGGGTGAAGATTATGATCAGACCATTGGTGGGATGACATTGAGA 1500
QY 423 pAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAsple-- 442

Db 1499 TAGCCTCCAAAGCCCACTTATTTCTCGTCAAGCGACAAGCGTGGAGGAAGAGATCGC 1440
QY 443 -ValHisIleGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyG 462
Db 1439 TGCACCTCATGGCAGTATATATGGTGTCTGGGA-----AGAGTAGTAGTCTCATGCA 1386
QY 462 uGlyLysAspGlyValSerSerThrAspIleGlyGlyTrpGlnLeuAlaTrpLysTr 482
Db 1385 GGGCGGGGAGCAGTAAGCAGCATGGCATTGTTGGGGATGGCAGTTGGCTTGGAAATG 1326
QY 482 pSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyriLeuHi 502
Db 1325 GACTGAGAGAGAGGTGCAGATGCGGAAAGAGGTGGCTTCCACAGTATCTACTTGTGA 1266
QY 502 sGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAs 522
Db 1265 TGAAGAGGTGTGACAGGTGATCGCAGGGCTCTATCTGTCTTCTCTCTCTCTCTCTCT 1209
QY 522 pValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPh 542
Db 1208 TGTTCCTCTGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1149
QY 542 eSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluVa 562
Db 1148 CTCIAGAAGATTGAGGACACGCTTCTGCGCCCTGCTATGGTGATCCATCTCAGGC 1089
QY 562 lAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLe 582
Db 1088 AGTTGCTAAAGTCCAAAATGGGCAGACTTATTTCGAACTCGAGTGAAGCATGCTCTGT 1029
QY 582 uValGlyValGlyLeuGlnIleLeuGlnPheAlaGlyLeuAsnGlyValLeuTyriTy 602
Db 1028 TGTGGCATAGGATACAAATCTTGCAACAGTTTCTGGCATTAATGGAGTCTGTACTA 969
QY 602 rThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSe 622
Db 968 CACTCCACAATTTCTTGAGCAAGCTGGTGTGGTGTCTTCTTCTGCAAAACATTGGACTTAG 909
QY 622 rSerAlaSerAlaSerIleLeuIleSerSerLeuThrLeuLeuLeuLeuProCysIl 642
Db 908 CTCCTCATCTCATCTATTTATTAGCGGACTGACACCTTGTCTGTCTTCTTCTTCTTCT 849
QY 642 eGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyTh 662
Db 848 TGGTATTCTATGAGCTCATGGATATGCTGGAAGAAGGTTTCTTCTCTTCTTCTTCT 789
QY 662 eProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyTh 682
Db 788 CCTATCTCTGATAGTACACTAGCTATCTTGTATCTGTGTAATATCTGATGTGGGAC 729
QY 682 rLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyriPheCysCysPheValMe 702
Db 728 CATGTTTCATGCTCTCATGTCACAGTCAGTGTCTACTCTACTCTCTCTCTCTCTCT 669
QY 702 tGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGl 722
Db 668 GGGGTTCGGGCTATTCCAAACATTTCTGTGACAGATTTTCCCGACACCGTTCTGTGG 609
QY 722 yLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyriSe 742
Db 608 CATCTGCATAGCCATCTGTGCGCTAACATTCGATCGGTGATATCATTTGACATACAC 549
QY 742 rLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyriAlaValVa 762
Db 548 CTTCCCGGTGATGCTCAACGCCATTGGACTCGCTGGAGTGTGGAACTCTACCGCATGTG 489
QY 762 lCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMetProIle 782
Db 488 CTGCATCTGCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 429
QY 782 uGluValIleThrGluPhePheAlaValIleAlaLysGlnAla 796

56	MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp	75
Qy		
175	ATGAAAGTGGCGTCTTGTGTTATTTGCCGCTTCCATTCGTATATTTCTCCAAGGATGG	234
Db		
76	AspAsnAlaThrIleAlaAlaValLeuTyrIleIysIysGluPheGlnLeuGlnAsn	95
Qy		
235	GATAATGCTACATCGCCGGCGCTAATGGTTACATTAAGAAAGACCTTGCTTTGGGAACA	294
Db		
96	GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr	115
Qy		
295	-----ACTATGGAAGGCTTGCTGGTGGCATGTCCCTGATTTGGAGCACGGTAATCACCC	348
Db		
116	ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer	135
Qy		
349	ACATGCTCTGTGCTTAGCGGATGTGCTGGTGGCGACCCATGATGATATCTCATCT	408
Db		
136	IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu	155
Qy		
409	GTGCTCTATTCTTGGGTGGTTTGGTGATGCTGTGTGTCCTCCCAATGTGTATGTGTGTGC	468
Db		
156	LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr	175
Qy		
469	TGCGCGAGGCTACTGTGATGGATTTGGGATTTGGCCTTCTGTGACTCTGTCCGGGTCTAT	528
Db		
176	IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer	195
Qy		
529	ATATCTGAAACGGCGCGCTCTGAATAAGGGGTGCTTGAATACGTTCTCTCATGTTCACT	588
Db		
196	GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer	215
Qy		
589	GGCTCTGGAGGAATGTTTTTGTCTGCTACTGTATGGTTTTTGGCATGTCAATTGAGTCCCGCG	648
Db		
216	ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu	235
Qy		
649	CCTAGCTGGAGGCTCATGCTTGGGGTTCGTGTATTCCTTCTCTTGTATTTTGCATTTG	708
Db		
236	ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerIysGlyArgMetAlaGlu	255
Qy		

709	Db	ACCAATTTTTTCTTGCCCGAGTCTCTCGGTGCTGGTCAGCAAAAGGAGATGCTCGAG	768
256	Qy	AlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeu	275
769	Db	GCTAAGAAGAGTCTCCAAAGATTGCGCGAAGGAGGATGTGTCAAGCGAGATGGCATTTG	828
276	Qy	LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro	295
829	Db	CTGGTTCAAGAGTCTCGGATTTGGGGTGATACATCTATCCAAAGAGTACATAATTGCCCT	888
296	Qy	AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr	315
889	Db	GCTGACGATGTGCTGATGGTCATGTAACATGCACACAGAGAAGATAAAATTCGATTATAT	948
316	Qy	GlyProGluGluGlnSerThrPheAlaArgProSerLysGlyProIleMetLeuGly	335
949	Db	GGATCCCAAGCAGCGCTTCTTGGTTATCAAAACCTGTCACTGGACAGAGTCTTATGGCG	1008
336	Qy	SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet	355
1009	Db	-----CTTCGGCTCACACCATGAAGCATCATCAACCAAGACATGCCCTCATG	1056
356	Qy	AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly-----	373
1057	Db	GATCCTCTGGTACACTGTTTGGTAGCATTCATGAAGAGCTCCCGAGACAGGACGAAGA	1116
374	Qy	GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln	393
1117	Db	GGAAGCATCGAAGCACTCTGTTTCCAAATTTTGGAAGCATGTTTCAGCACTGCTGACCGC	1176
394	Qy	HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla	413
1177	Db	CATGCTAAATTCGAACAATGGGATGAAGAAGCTTACAAAGGGAGCGTGAAGACTACATG	1236
414	Qy	SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln	433
1237	Db	TCAGATGCACACCGTGGGACCTCCGATGATAAATTTGCAGATCCTTTATCTCAGCCCAA	1296
434	Qy	AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeu---	452
1297	Db	ACMACAAGCCTTGA---AAAGACTTACCTCTCTCTCTCCATGCCAGTACTCTTGGC	1353
453	Qy	SerMetArgArgGlnThrLeuLeuGlyGluGly---GlyAspGlyValSerSerThrAsp	471
1354	Db	AGCATGAGGCGTCACAGTAGTCTCATGCAAGGCTCAGGTGAGCAAGGTGGTAGTACAGT	1413
472	Qy	IleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyValAsnGlyArg	491
1414	Db	ATTGGTGGTGGTGGCACTGCATGGAATGGACTGATAA---GGTGAGGATGAAAA	1470
492	Qy	LysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArg	511
1471	Db	CAACAAGGAGGGTTTAAAGGATTTATTACATGAGGAGGAGTTTCTGCACTCTCGTCT	1530
512	Qy	GlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHis	531
1531	Db	GGATCCATTGTATCGATTCCCGGTGAAGC-----GAATTTGTCACAG	1572
532	Qy	AlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMet	551
1573	Db	GCTGCTGCTGTTGAAGCCAAACCCGCTCTTTACTCCAAGGAGCTTATTGATGGACACCCA	1632
552	Qy	SerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAsp	571
1633	Db	GTTGGGCTGCAATGGTTTCAACCATCTCGACAGAGCTTCAAGGGGCCAAGTTGGAAGCT	1692
572	Qy	LeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln	591
1693	Db	CTTCTTGAAACCAAGGGTTAAGCATGCATTTGGTTGGAGTTGGAAATACAAATACTTCAG	1752
592	Qy	GlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGly	611
1753	Db	CAGTTTTTCAGGATTAATGGGGTCTATTATACACACCTCAATCTTCAAGAGCCCGT	1812

612 ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer 631
1813 GTTGAAGTCTCTTTTCAGATATAGGATTTGGCTCAGAGTCGGCATCATCTTATCAGT 1872
632 SerLeuThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu 651
1873 GCTTTCAACACCTTCTTATGCTTCCCTGTATAGGCTAGCAGTGAAGCTCATGGATGT 1932
652 SerGlyArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671
1933 TCAGGACAGAGCAGTGTCTACTTACATCAATCCCGTGTGATGTGTCACTCAATATT 1992
672 LeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691
1993 TTGGTCATTTGAAGCCTGGTAAATTTTGGCAATCTGCCCATCGCAGCAATCTCAACAGTA 2052
692 SerValIleValTyrPheCysPheValMetGlyPheGlyProIleProAsnIleLeu 711
2053 TCGGTGTGGTATTATCTCTGCTGCTTGTGATGGTTATGGACCAATTCACCAATCCTT 2112
712 CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731
2113 TGCTCAGAGATTTTCCCACTAGGCTGGTGGCTCTGCAATGTCTATCTGTGATTAAGT 2172
732 PheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGly 751
2173 TTCTGATTTGGAGACATCATCATCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2232
752 LeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPhe 771
2233 CTGTGTGTGTATTCGGCAATTTAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2292
772 LeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaVal 791
2293 TTGAAGTTCAGAAACAAAGAGGATGCCCTTGAAGTCTATCTGAAATCTTTCTGTT 2352
792 GlyAlaLysGlnAlaAlaA 798
2353 GGAGCAAGCAGGCTGCTTCT 2373

RESULT 8

US-10-051-909-7
; Sequence 7, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: B01163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: Glycine max
US-10-051-909-7

Alignment Scores:

Pred. No.:	7,15e-255	Length:	2601
Score:	2674.00	Matches:	522
Percent Similarity:	82.20%	Conservative:	92
Best Local Similarity:	69.88%	Mismatches:	115
Query Match:	65.62%	Indels:	18
DB:	13	Gaps:	8

US-10-051-909-32 (1-800) x US-10-051-909-7 (1-2601)
QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr 75
Db 175 ATGAAAGTGGCGTCTTGTGCTATTTGGCTTCAATTTGCTATTTCTTCCAGGATGG 234
QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysGlyGluPheGlnLeuGlnAsn 95
Db 235 GATAATGCTACCATCGCGGGGCTTAATGGTTATCAATTAAGAAAGACCTTGTCTTTGGACA 294
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db 295 -----ACTATGAAAGGCTTGTGGTGGCATGTCTCTGATTTGGAGCAACGGTAATCACC 348
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgProMetLeuLeuLeuSerSer 135
Db 349 ACATGCTCTGTCTCTATAGCGATTTGGCTCGTGGCGACCCATGATGATAATCTCATCT 408
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu 155
Db 409 GTGCTCTATTTCTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 468
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
Db 469 TTGGCGAGGCTACTGTGATTTGGGATTTGGCTTGTGACTCTTGTCTCCGGTCTAT 528
QY 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
Db 529 ATATCTGAAACGGCGCGCTCTGAAATAAGGGGCTGTTGAAATACGCTTCTCTAGTTCAGT 588
QY 196 GlySerGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
Db 589 GGCCTCTGGAGGAATGTTTTGTCTGTATGTTGGTGGTGGTGGTGGTGGTGGTGGTGG 648
QY 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
Db 649 CTAAGCTGGAGGCTACGCTTGGGTGCTGTCTATTTCTCTCTCTGTATTTTGGATTG 708
QY 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255
Db 709 ACCATTTTTTCTTGGCCGAGTCTCTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 768
QY 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeu 275
Db 769 GCTAAGAAAGTGTCTCCAAAGATTGGCGGAGGAGGAGTGTCTCAGCGGAGATGGCATTTG 828
QY 276 LeuLeuGluGlyLeuGluValGlyAspThrSerIleGluGluTyrIleleGlyPro 295
Db 829 CTGTTTCAAGGCTCTCGGCTTGGGGTGTATATCTATCTATCTCTCTCTCTCTCTCTCT 888
QY 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315
Db 889 GCTGACGATGTGCTGATGCTCATGACATGCAACAGAAAGATAAATTCGATTATAT 948
QY 316 GlyProGluGluGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335
Db 949 GGATCCCAAGCAGGCTTTCTTGTGTTATCAAAACCTGTCTACTGGACAGAGTTCTATTGGC 1008
QY 336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355
Db 1009 -----CTGGCTCACCATGGAGATCATCAACCAAGCATGCCCTCATG 1056
QY 356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly----- 373
Db 1057 GATCCTCTGGTGACACTGTTTGGTAGCATTCATGAGAGCTCCCGGACAGCAGCAAGA 1116
QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db 1117 GGAAGCATGGAAGCACTCTGTTTCCAAATTTTGGAGCATGTTTCAGCTACTGCTGAGCGG 1176
QY 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluTyrAla 413
Db 1177 CATGCTAAATTCACAATGGGATGAAGAAAGCTTACAAAGGCAACGCTGAGGACTACATG 1236

196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
 421 GGCCTCTGGAGGAGTGTCTTATCTTACGTATGTTTTCGGAATGCTGTTGATCCATCA 480
 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
 481 CTTAGCTGAGAGTGTATGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 540
 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255
 541 ACGGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 600
 256 AlaLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeu 275
 601 GCRAAGCGGTTCTTACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 276 LeuLeuGluGlyLeuGluValGlyAspThrSerIleGluGluTrpIleIleGlyPro 295
 661 TTGGTTGAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 720
 296 AlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315
 721 CGGATGAGTTACTGATGATCATGATATAGCTGTGGATAGGATCAATTAATTTATAT 780
 316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335
 781 GGTGCAGAGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 336 SerValLeuSerLeuAlaSerArgHisGlySer---MetValAsnGlnSerValProLeu 354
 840 ACATCTAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 888
 355 MetAspProIleValThrLeuPheGlySerValHisGlnAsnMetProGlnAlaGlyGly 374
 889 ATTGATCTCTTGTCCACACTGTTGGAGCGTTCACGAGAGATGCGGACACT---GGA 945
 375 SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHis 394
 946 ACATCTAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1005
 395 AlaLysAsnGluGlnTrpAspGluAlaAsnLeuHisArgAspAspGluTrpAlaSer 414
 1006 CCAAGACATGAAGTGGGATCAAGAGATCTTGTGGAGAGGTGAGGATATCCATCC 1065
 415 AspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla 434
 1066 GAC---CATGGAGATGATCTGAGATGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1122
 435 ThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGlySerAlaLeuSerMet 454
 1123 ACAAGCATGGAG---AAAGACATGCTCCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1179
 455 ArgArgGlnThrLeuLeuGlyGlyGlyAspGlyValSerSerThrAspIleGlyGly 474
 1180 CATGGAATGCAAGTCAAGGAGCTCAAGGGAAGAGCGGTAGTAGTGGGATGGAGGT 1239
 475 GlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGluAsnGlyArgLysGluGly 494
 1240 GGATGGCAAGTGGCATGGAATGGAGAGAGAGATGAATGCGGACAGAAAGAA--- 1296
 495 GlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgGlySerIle 514
 1297 -----GAGGTTTCCGAGATCTCGAGTGGCTCAAT 1329
 515 ValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaAla 534
 1330 GTTTCATTGCTGCTGGTGGTGGATGGAACCGGTGAG---GCAGATTTTGTACAGGCTGCT 1386
 535 LeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAla 554
 1387 TTGGTTACCAACCAAGCTCTTTATTCCAAGACCTCTCAAGAGAACATCAATTTGCTCT 1446
 555 AlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGlu 574

1447 GCTATGGTACATCCATCCGAA---ACAACTAAAGGGTCAATTTGGCATGATCTTCATGAT 1503
 575 ProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPheAla 594
 1504 CTTGAGTCAAGCGTGCATTTAGTCTGAGGATGGACTTCAATATCTTACAGAGTTCTCA 1563
 595 GlyIleAsnGlyValLeuTyrTrpProGlnIleLeuGluGlnAlaGlyValAlaVal 614
 1564 GGCATCAACGGAGTCTTTTACTACACACCGCAATCTTGGAGCGGGTGTGGGATC 1623
 615 IleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerLeuThr 634
 1624 CTACTATCAACATGGGATTTAGTCTTCTCAGCATCTTACTTATTAAGTCAATGACA 1683
 635 ThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg 654
 1684 ACCTTTGTGATTTACCTGCAATAGCTTGTCAATGAGGCTCATGGATCTTTCTGGTCA 1743
 655 ArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValVal 674
 1744 AGGACCTTCTTCTTCCACCATACCAATCTCTGATAGCATCTCTATTTGGTTTAGTAA 1803
 675 SerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValle 694
 1804 TCAAACTTGTTCATCATGACAGCATGTCACGCGTCTTATCAACCGTAAGCGTTGTG 1863
 695 ValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlu 714
 1864 CTCATCTTCTCTTCTTCTGATGCGGTTTGGTCTCTCTCTCAAAACATCTCTGTTCA 1923
 715 IlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIle 734
 1924 ATTTTCCAACTCGAGTCCGCGAATCTGCATCGCCATCTCGGCACTTCACCTCTCTG 1983
 735 GlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGly 754
 1984 TGTGACATPANTGTCATCTTACGCTCTCCCGTGTCTCTCAAAATCCATGAGCTAGT 2043
 755 ValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysVal 774
 2044 GTGTTTGAATGTACGCAATCGTATGTGATTTTCATGGTCTTGTGTTCAATTAAGTC 2103
 775 ProGluThrLysGlyMetProLeuGluValIleThrGluPheAlaValGlyAlaLys 794
 2104 CCGAAACTAAAGGATGCCACTTGAAGTCAATCAGAGTCTTTCTGTTGGAGCTAGA 2163
 795 GlnAlaAlaAla 798
 2164 CAAGCTGAAGCT 2175

RESULT 10
 US-09-938-842A-1315
 ; Sequence 1315, Application US/09938842A
 ; Publication No. US20040009476A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Krebs, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE OF INVENTION: SAME, AND METHODS OF USE
 ; FILE REFERENCE: SCRIP1300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; PRIOR FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 1315

829	Db	AGCACTATGAGTGTGTTTGTCTCGCCATGGAAGTACAAATGACGAGGAGCGCAAGGCTCATTTG	898
355	Qy	MetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGly	374
889	Db	ATTGATCTCTGTGCACACTGTTTGGAGCGTTTCAGAGAAGATGCCGGACACT--GGG	945
375	Qy	SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHis	394
946	Db	AGCATGAGGAGTGCCTTGTTCACACTTTTGGGAGTATGTTCACTGTTCAGTGTGAGGGAATCAA	1005
395	Qy	AlaLysAsnGluGlnThrAspGluGluAsnLeuHisArgAspAspGluGluThrAlaSer	414
1006	Db	CCAAGAATCAAGAATTTGGGATGAAGAATCTTGTGGAGAAGGTGAGGATATCATCCTCC	1065
415	Qy	AspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla	434
1066	Db	GAC--CATGGAGATGATCTGAAGATGATCTTCATCTCCGTGTGATCTCACGTCAAACG	1122
435	Qy	ThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGlySerAlaLeuSerMet	454
1123	Db	ACAAGCATGGAG--AAAGACATGCCTCCACATGCCTCATCGGAACCTTTTCTACCTTCAGA	1179
455	Qy	ArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGly	474
1180	Db	CATGGAAGTCAAGTCGAGGAGCTCAAGCGGAAGGACGGGTAGTATGCGGATTCGAGGT	1239
475	Qy	GlyTyrGlnLeuAlaTyrIlePheSerGluLysGluGlyGluAsnGlyArgLysGluGly	494
1240	Db	GGATGCAAGTCGATCGAAATCGGACGGAAGAGAGATGAATTCCTCGGACAGAAAGAA--	1296
495	Qy	GlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgGlySerIle	514
1297	Db	-----GAAAGTTTCCGAGGATCTCGACGTGGCTCAATTT	1329
515	Qy	ValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaAla	534
1330	Db	GTATTCAATTCCTGGTGGTATGGAACCGGTGAG--GCAGATTTTGTACAAAGCGTCTGCT	1386
535	Qy	LeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAla	554
1387	Db	TTGGTTAGCCACACGACTCTTTTTCCAAAGACCTTCTCAAAGAACATPACAAATTTGGTCT	1446
555	Qy	AlaMetValHisProSerGluValAlaAlaLysGlySerArgTyrPheAspLeuPheGlu	574
1447	Db	GCTATGGTATCATCCATCCGAA--ACAACATAAGGGTCAATTTGGCATGATCTTCATCAT	1503
575	Qy	ProGlyValArgAlaAlaLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAla	594
1504	Db	CCTGGAGTCAAGCGTGCAATTAGTCGTGGAGTTGGACTTCAAATACTTCACACAGTCTCA	1563
595	Qy	GlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaVal	614
1564	Db	GGCATCAACGGAGTCTTTACTACACACCGCAATCTTTGAGCAGCGCGGTTCGGGATC	1623
615	Qy	IleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThr	634
1624	Db	CTACTATCGAACATGGGATAGTTTCTTCCTCAGCATCTTACTTATAAGTGCATTCACA	1683
635	Qy	ThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg	654
1684	Db	ACCTTTGTGATGTTACCTTGCATAGCTGTTGCATAGGCTCATGGATCTTCTGGTTCGA	1743
655	Qy	ArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValVal	674
1744	Db	AGGACCTTGTCTCTCACCACCATACCAATCCTGATAGCATCTCTATTGGTTTTAGTAATC	1803
675	Qy	SerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIle	694
1804	Db	TCAATCTTGTTCACTGAACACGATTTGTGCACCGGTCTTATCAACCGTAAGCGTTGTG	1863
695	Qy	ValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlu	714
1864	Db	CTCTACTTCTGCTCTCTCGTATCGGTTTCGGTCTGCTGCTCCTGCTCCTGCTGCTGCTG	1923

QY 296 AlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315
Db : : : : :
796 GCCAATGAATTCAGT-----GATCAGAGCAGATCAATTTGTAT 834
QY 316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335
Db : : : : :
835 GGGACAGCAGAGGTCAATCCTCGAATGCCAAACCTGTCACTGGCAAAAGTTCCATTGGC 894
QY 336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnGlnSerValProLeuMet 355
Db : : : : :
895 -----CTTGATCTCGGAGGAGGAGCATGCCAATCAAGTGTCT--CTAGTG 939
QY 356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySer 375
Db : : : : :
940 GACCTCTAGTGAAGCTCTTTGGTAGTGTCCATGAGAAGCTCCCAACAGCA----- 993
QY 376 MetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHisAla 395
Db : : : : :
994 -----AGCAACCTTTCCACACTTTGGGAAGCATGTTTAGTGTGGGGGAATCAACCT 1047
QY 396 LysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTrpAlaSerAsp 415
Db : : : : :
1048 AGCAATCAAGATTGGGATGAAGAAAGCATTCGCCAGAGGAGGATGATTATGTATCTGAT 1107
QY 416 GlyAlaGly---GlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla 434
Db : : : : :
1108 GCTGCTGACACCGACGATCTGATGACAAATTTGCAGAGTCCATGATCTCACGTCACGCA 1167
QY 435 ThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGlySerAlaLeuSerMet 454
Db : : : : :
1168 ACAAGTCGGAA---AGGGACATGCTCTGCTGCCAGGGA-----AGCATG 1212
QY 455 ArgArgGlnThrLeuLeuGlyGlyGlyValSerSerThrAspIleGlyGly 474
Db : : : : :
1213 AGCAAGGTAGTCTTTTG-----CRAAGAGAACCTGCTCGAAACCTCAGGGATTTGGTGT 1266
QY 475 GlyTrpGlnLeuAlaTyrLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGly 494
Db : : : : :
1267 GGTGGCAGCTAGCATGGAAATGTCTGAACAGAGGTGTT----- 1308
QY 495 GlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIle 514
Db : : : : :
1309 ---TTCAAGAGAATATATTTCACCAAGAGGTGCTCTGGATCTAGACGTGGGTCCATC 1365
QY 515 ValSerLeuProGly---GlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533
Db : : : : :
1366 ATTTCACTTCAGGCTGTGATGCGCGACACTAACAGATGGTGAGATTGTCAGGCTGCC 1425
QY 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553
Db : : : : :
1426 GCTCTAGTGAATCAATCAGCCCTTTACACAGAGGCTTATGATCAGCAGCCAGTTGGA 1485
QY 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe 573
Db : : : : :
1486 CCAGCCATGATTCATCCATCCCAACAGCTGCCAAAGGGCGGAGTTGGAGTGACCTTTTA 1545
QY 574 GluProGlyValArgArgAlaLeuValGlyValGlyIleGlnIleLeuGlnPhe 593
Db : : : : :
1546 GAGCCTGGAGTGAAGCATGACATGATTTGGGGGTGGGATTCAAATTTCTTCAACAGTTC 1605
QY 594 AlaGlyIleAsnGlyValLeuTyrTrpProGlnIleLeuGlnAlaGlyValAla 613
Db : : : : :
1606 TCTGGTATAAATGGGTGCTCTACTATACACCTCAAATTTCTTGCAGCAGGAGGTGTGT 1665
QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
Db : : : : :
1666 TATCTTTTCAAACCTGGGCGCTGGTCTTCTCTGCGTCTTTCTTATCAGCAGCGTG 1725
QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
Db : : : : :
1726 ACAACCTTTGTATGTTTACCTGTATAGCTGTGCCATGAGGCTCATGGATATATCTGGC 1785

QY 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
Db : : : : :
1786 AGAAGACCCCTGCTGCTGACTACAAATCCCGCTTTGATAGTATCACTTCTCATATTAGTA 1845
QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
Db : : : : :
1846 ATAGAGAGTCTTGTGGAATTTGGAGTAGTACTATAAATGCTTTTATCTCACTTCGAGTGT 1905
QY 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
Db : : : : :
1906 ATGTCTACTTCTGTGTTTCTGTCATGGCTTTGGCCCAATTCCTAATATCTTTGTTCA 1965
QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
Db : : : : :
1966 GAGATCTTCCCTACTCAGTGTGCTGCTATGCATAGTATATGTGCTCTTACATTTTGG 2025
QY 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
Db : : : : :
2026 ATATGTGACATCATAGTACCTACTCGCTCCAGTTATGCTTAATCTGTGGCCTTGGT 2085
QY 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773
Db : : : : :
2086 GGTGTCTTTGGTATGATGCGGTTGTGTGCATCATAGCTGGGTGTTGTCTTCTTAAAA 2145
QY 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPheAlaValGlyAla 793
Db : : : : :
2146 GTTCAGAAACCAAGGCGATGCCACTGGAAGTTATCATGAGTCTTCTCTGTCGGGCA 2205
QY 794 LysGlnAla 796
Db : : : : :
2206 AAGCAAGCT 2214
RESULT 13
US-10-424-599-108435
; Sequence 108435, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 108435
; LENGTH: 2573
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_68931C.1
US-10-424-599-108435
Alignment Scores:
Pred. No.: 5,866-236 Length: 2573
Score: 2483.50 Matches: 505
Percent Similarity: 76.69% Conservative: 84
Best Local Similarity: 65.76% Mismatches: 121
Query Match: 60.94% Indels: 58
DB: 16 Gaps: 12
US-10-051-909-32 (1-800) x US-10-424-599-108435 (1-2573)

QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuGlnGlyTrp 75
Db : : : : :
37 ATGAGCGGAGCTGTTCTTGTGTGCTGCTGCTATCGGCAACTTCTGCAAGGATGG 96
QY 76 AspAsnAlaThrIle----- 80
Db : : : : :
97 GATATGCCACCATTCGAGGTCATATTTTATTTCGTTAGCAATTAAGATAACATTCAA 156

Qy 81 -----AlaAlaAlaValLeuTyrIleLysLysGluPhe 91
Db 157 CCAAGTCTAATTGGCTTTCTATGCTATGAGGATCTATTGTTACATAAAGAGGAGTTT 216
Qy 92 GlnLeuGlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAla 111
Db 217 AATTTCACAAAGTGAACCCACCAATAGAAAGTCTAATTGTGGCCATGTCACCTATTGGAGCC 276
Qy 112 ThrIleValThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeu 131
Db 277 ACTGTCGTACGACGCTCTGACCCCTATACAGACTTGTAGCCCGGCGTCTTAGTTG 336
Qy 132 IleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTyrPheSerProLeuVal 151
Db 337 ATCATGCTCTATCTCTTATTGTTGTTCTCTGCTATGCTATGCTATGCTCAATGTT 396
Qy 152 TyrValLeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeu 171
Db 397 TATATCTCTCTTTGGAGGCTTTAGATGGATTAGGCAATTGGTTGGGGTCCACTTG 456
Qy 172 ValProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeu 191
Db 457 GTACCTCTTTACATATCCGAGACAGCTCCACCTGAGATTAGGGGATTACTAAATCTCT 516
Qy 192 ProGlnPheSerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSer 211
Db 517 CCGCAGTCACTGGTCTCTGCGAATGTTCTTTCTCTACTGATGCTCTTCTGCTATGCTA 576
Qy 212 LeuSerProSerProAspTPrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhe 231
Db 577 CTCACCAAGGCCCAAAATGAGACTCATGTTGGGTGTTCTTTCAATTCCTCTCTCAT 636
Qy 232 PhePheGlyLeuThrIlePheTyrLeuProGluSerProArgThrLeuValSerLysGly 251
Db 637 TATTTTGGCACTACGCTATTTCTTTCTGCTGAAATCTCCAGATGGGTGTTGTC---AAAGGC 693
Qy 252 ArgMetAlaGluAlaLysValLeuGlnLysLeuArgGlyLysAspValSerGly 271
Db 694 CGATGCTAGAGGCCAAGAGGTTTTCACAGACTTCGGCGCAGACAGATGTCGTGT 753
Qy 272 GluLeuSerLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyr 291
Db 754 GAGATGCTTTACTAGTTAGGGTCTTGGAGTTGGCGGTGATACCCGCTATAGAAGAGTAC 813
Qy 292 IleIleGlyProAla-ThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGly 311
Db 814 ATAATTGGTCCAGCTCAATGAATTCAGT-----GAGCGACAGCA 852
Qy 311 nIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyPr 331
Db 853 GATCAAAATTGTATGGGACACAGAGGTGATCTCGAATGCTTAAACCTGTCACTGGACA 912
Qy 331 oIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSe 351
Db 913 AAGTTCATTTGGT-----CTTGATCTCGAAGGGAAGCATGGCAATCAAAG 960
Qy 351 rValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProgl 371
Db 961 T---GGTCTAGTGACCCCTCTAGTGAACCTTTTGGTAGTGTCCATGAGAAGCTCCACA 1017
Qy 371 nAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValTh 391
Db 1018 AACACGGA-----ASGCACATTTTCCACACTTTGGAAGCATGTTAGTGTGG 1065
Qy 391 rAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspGluGly 411
Db 1066 GGGGAATCAACCTAGGAATGAGATTGGGATGAAGAAAGCATTCGCCAGAGAGGGAGATGA 1125
Qy 411 uTyrAlaSerAspGlyAlaGly---GlyAspTyrGluAspAsnLeuHisSerProLeuLe 430
Db 1126 TTATGTATCTGCTGCTCTGACACCCGACGATTTCTGATGACAAATTCAGAGTCCATTGAT 1185
Qy 430 uSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGlySe 450

Db 1186 CTCACGTCACAGCAACAAGTCGGAA---AGGGACATGCCTGCTCCTGCCCCAGGGA----- 1237
Qy 450 rAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerTh 470
Db 1238 -----AGCATGAGGCAAGGTAGTCTTTTG-----CAAGGAGAACCTGCTGAAACTC 1284
Qy 470 rAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGly 490
Db 1285 AGGANTTGGTGGTGGATGGCAGCTAGCATGGAATGGTCTGAAACAGAAAGTCTT----- 1339
Qy 490 yArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerAr 510
Db 1340 -----TTCAGAGAATATATTTGCAACAAGAAGGTGCTCTGGAICTAG 1383
Qy 510 gArgGlySerIleValSerLeuProGly---GlyGlyAspValPheGluGlySerGluPh 529
Db 1384 ACGTGGGTCCATCATTTTCACTTCCAGGCTGTGATGCCGCCACACTAACAGATGTTGAGAT 1443
Qy 529 eValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluPr 549
Db 1444 TGTCCAGCTGCCCTCTAGTAGTCAATCAGCCCTTTACACAAGAGGCTTATGCAATCA 1503
Qy 549 oArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTr 569
Db 1504 GCAGCCAGTTGGACCGCATGATTCATCCATCCAAACAGCTGCCAAAGGGCGGAGTTG 1563
Qy 569 pLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnI 589
Db 1564 GAGTGACCTTTAGAGCTTGGAGCTGAGTGAAGCATGCACTGATTGTGGGGTGGGAGTTCAAT 1623
Qy 589 eLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluG 609
Db 1624 TCTTCAACAGTCTCTGGTATTAATGGGTGCTCTACTATACACTCAATCTTGTAGCA 1683
Qy 609 nAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLe 629
Db 1684 GGCAGGTGTTGGTTATCTTCTTTCAACCTCGGGCTCGGTTCTACTTCTCGCTTTTCT 1743
Qy 629 uIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMe 649
Db 1744 TATCAGCAGCTGACAACTTGTGATGTTACCTGTATAGCTGTGCCATGAGGCTCAT 1803
Qy 649 tAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLe 669
Db 1804 GGATATATCTGGCAGAAAGAACCTTGTCTGCTGACTACAAATCCCGCTCTTGTATGATCACT 1863
Qy 669 uValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSe 689
Db 1864 TCTCATATTAGTAATAGGAGTCTGTGGAATTGGAGACTATATAAATGCTTTTATCTC 1923
Qy 689 rThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAs 709
Db 1924 AACITCGAGTGTATTGTCTACTTCTGTTCTGCTCATGGGCTTGGGCAATTCCTAA 1983
Qy 709 nIleLeuCysAlaGluLeuPheProThrArgValArgGlyLeuCysIleAlaIleCysAl 729
Db 1984 TATCCTTGTTCAGAGATCTTCCCTACTCGAGTTCGGTCTATGACATGATATATGTC 2043
Qy 729 aPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAl 749
Db 2044 TCTTACATTTGGATATGTGACATCATGACCTACTCGCTCCACGTTATGCTTAATTC 2103
Qy 749 aIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPh 769
Db 2104 TGTGGGCTTGGTGGTCTTCTTGTGATGTCGGTGTGTGTCATCATGCTGGGTGTT 2163
Qy 769 eValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePh 789
Db 2164 TGTCTTCTTAAAGGTTCCAGAACCAAGGGCATGCCACTGGAAGTTATCATTTGAGTTCTT 2223
Qy 789 eAlaValGlyAlaLysGlnAla 796
Db 796 eAlaValGlyAlaLysGlnAla 796

Db 2224 CTCTGTCGGGGCAAAAGCAAGCT 2245

RESULT 14

US-10-425-114-29146

Sequence 29146, Application US/10425114

Publication No. US20040034988A1

GENERAL INFORMATION:

APPLICANT: Liu, Jingdong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E

APPLICANT: Tabaska, Jack E

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53313)B

CURRENT APPLICATION NUMBER: US/10/425,114

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128

SEQ ID NO 29146

LENGTH: 2271

TYPE: DNA

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Clone ID: LIB83-005-C10_FLI

US-10-425-114-29146

Alignment Scores:

Pred. No.: 5,7e-232 Length: 2271

Score: 2442.50 Matches: 472

Percent Similarity: 83.66% Conservative: 76

Best Local Similarity: 72.06% Mismatches: 100

Query Match: 59.94% Indels: 7

DB: 16 Gaps: 6

US-10-051-909-32 (1-800) x US-10-425-114-29146 (1-2271)

Qy 143 LeuileMetLeuTrpSerProAnValTyValLeuLeuLeuAlaAargPheValAspGly 162

Db 3 CTGGTGATGCTCTGGGGCCAGTGATACATCTTGCTCTCGCAAGCTCATTTGTTGG 62

Qy 163 PheGlyIleGlyLeuAlaValThrLeuValProLeuTyTrIleSerGluIleAlaProSer 182

Db 63 TTGGGTATCGGTTGGCGGTACACATGTCTCTCTACATCTCGAGACGCGCCGACAC 122

Qy 183 GluileAargGlyLeuLeuAnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeu 202

Db 123 GACATTCGTGGGTTGTGAACACGTCGCGCAGTTCAGTGGGTGAGAGGATGTTCTCTC 182

Qy 203 SerTyrcysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMetLeu 222

Db 183 TCCTACTGTCATGTTGGGATGTCCTCATGCCCAACCTGATTTGGAGGCTCATGCTT 242

Qy 223 GlyValLeuAlaIleProSerIleuPhePheGlyLeuThrIlePheTyTrIleProGlu 242

Db 243 GGAGTTCGTGATCCGCTGCTTACTTTGGACTGACTGCTCTTCTACTTGCCTGAA 302

Qy 243 SerProArgTrpLeuValSerGlyAargMetAlaGluAlaLysValLeuGlnLys 262

Db 303 TCACCAAGTGGCTTGTGACAAAGAGGATGGCTGAGCGCGAAGAGAGTGTGCAAGG 362

Qy 263 LeuArgGlyLysAspValSerGlyGluLeuSerLeuLeuLeuGlyLeuGluVal 282

Db 363 CTGGGGGAAGAGATGCTCAGGGGAGATGGCTCTTCTAGTTGAAGGTTTGGGGGTC 422

Qy 283 GlyGlyAspThrSerIleGluTyTrIleIleGlyProAlaThrGluAlaAlaAspAsp 302

Db 423 GGTAAAGATACACGTATTGAAGAAATACATATTGGTCCCGATGATGAACCTTGTGATGAA 482

Qy 303 LeuValThrAspGlyAspLysGluGlnIleThrLeuTyTrGlyProGluGluGlyGlnSer 322

Db 483 GGGCTGGCTCCA---GATCCAGAGAGATCAAACTATATGGACCTGAAGAGGCGCTATCT 539


```

Db 1605 GACGTGGCGCAGCGGGCTCTCCAGCGCGTATAGTCTACTTCTGCTTCGTC 1664
Qy 702 MetGlyPheGlyProLeuProAsnLeuLeuCysAlaGluLeuPheProThrArgValArg 721
Db 1665 ATGGGGTTCGGGGCCGCCCAACATCTCTGGCGCAGAGATCTTCCCCACCGGTCGCG 1724
Qy 722 GlyLeuCysIleAlaIleCysAlaPheThrPheThrPheThrPheThrPheThrPheThr 741
Db 1725 GGTGTCGCAATCGCATCTCGCCCTGCGGCTTCTGGTTCGGCGACATCATCGTACGATAC 1784
Qy 742 SerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaVal 761
Db 1785 ACTCTCCCGCTGATGCTGAACGTCGTCGGCTCGCGGGCTCTTTGGGGTGTACGCCGTC 1844
Qy 762 ValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMetPro 781
Db 1845 GTGTGGCTCTCTAGCCCTCGCGTTCGGTTCATCAAGGTGCCCGAGACGAAGGCGATGCGT 1904
Qy 782 LeuGluValIleThrGluPhePheAlaValAlaValAlaValAlaValAlaValAlaVal 796
Db 1905 CTCAGGTCTATCACCAGTCTCTCTCGTTCGGTGGCGCAAGCAAGCC 1949

```

RESULT 15

```

US-09-938-842A-2254
; Sequence 2254, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Krepes, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2254
; LENGTH: 2205
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2254

```

```

Alignment Scores:
Pred. No.: 2,366-215 Length: 2205
Score: 2275.00 Matches: 462
Percent Similarity: 75.89% Conservative: 111
Best Local Similarity: 61.19% Mismatches: 146
Query Match: 55.83% Indels: 36
DB: 9 Gaps: 14

```

US-10-051-909-32 (1-800) x US-09-938-842A-2254 (1-2205)

```

Qy 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlyTyr 75
Db 1 ATGAAGGAGCGACTCTCGTCTGCTCGCGCCACATCGGCAATCTTACAGGATGG 60
Qy 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysGluPheGluLeuLysAsn 95
Db 61 GACATGCCACCATCTCTCGGAGCTATGGTTATATCAACAAGACTTGAATCTA----- 114
Qy 96 GluProThr---ValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleVal 114
Db 115 ---CCAACCTCTGTTCAAGGCTCTGTCGTTGCTATGTCATTGATCGGTGCAACGTCATC 171
Qy 115 ThrThrPheSerGlyProLeuSerAspSerIleGlyArgProMetLeuIleLeuSer 134

```

```

Db 172 ACGATTGCTCAGACCGATATCTGATGGCTCGGCGAGCGCCCATGCTCATTTATCA 231
Qy 135 SerIleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeu 154
Db 232 TCAGTTATGATTTCGCTCGCGTTCGATTAATGTTGTGGTCTCCCAATGTCATGTTCTG 291
Qy 155 LeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeu 174
Db 292 TGCCTTGTAGCTTCTTAATGGGTTTGGTCCCGGCTCGCGGTACACTTGCCCTGTT 351
Qy 175 TyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPhe 194
Db 352 TACATTTCTGAACCGCTCTCCGAGATCAGAGGACAGTTAAATACTCTCCCTCAGTTT 411
Qy 195 SerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerPro 214
Db 412 CTTGGCTCTGGTGAATGTTTGTTCATCTGATGTTTTCATGTTCTGCTGAGTGAC 471
Qy 215 SerProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGly 234
Db 472 TCCCTAGCTGGAGCGCATGCTCGGTGCTCTCGATCCCTCTCTCTTATTGTTT 531
Qy 235 LeuThrIlePheTyrLeuProGluSerProArgTyrLeuValSerLysGlyArgMetAla 254
Db 532 CTCACGGTGTATTTATTTGCCCGAGTCTCTCTGTTGGTGTAGTAAAGGAAGATGGAC 591
Qy 255 GluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSer 274
Db 592 GAGCTAAGCGAGTCTTTCACAGTATGTGGCAGAGAGATGTTACCGTAGATGGCT 651
Qy 275 LeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGlyIleLeuGly 294
Db 652 TTACTAGTTCAAGGACTAGATAGGAGGAGAAAAACAATGGAAGATCTCTTAGTAAC 711
Qy 295 ProAlaThrGluAlaAspAspLeuValThrAspGlyAspLysGlu---GlnIleThr 313
Db 712 TTGGAGGATCATGAGGTGATGATACACTTGAACCGTGTGATGAGGATGCAAAATGCGG 771
Qy 314 LeuTyrGlyProGluGluGlyGlnSerTyrIleAlaArgProSerLysGlyProIleMet 333
Db 772 CTTTATGGAACCCACGAGATCAATCGTACCTTGTAGACTGTCCAGACAA----- 825
Qy 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
Db 826 ---AATAGCTCCTTGGGCTACGCTCTCGCCACGGAAGCTTAGCAACCAAGCATGATC 882
Qy 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
Db 883 CTTAAAGATCCGCTCGTCAATCTTTTGGCAGTCTCCACGAGAAGATGCCAGAGCAGGC 942
Qy 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db 943 GGAACACTCGGAGTGGGATTTCCCTCATTTCCGGAAGCATGTTCACTACTACTGCGCAT 1002
Qy 394 -----HisAlaLysAsnGluGlnTyrAsp-----GluGluAsnLeuHisArgAsp 408
Db 1003 CGCCTCTACGTTAAACCGGCTCATTCGGAAAGGACATAGAGAGCCATTACACAAGAC 1062
Qy 409 AspGluGlyTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAsp-----AsnLeu 425
Db 1063 AATGATGACTATCGGACTGATGATGTCGGGGTGTGATGATGATGATCGGACCAACATTTG 1122
Qy 426 HisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHis 445
Db 1123 GTAGGCCCTTATATGTCGCCAGACACAGCATGGAC---AAGGATATGATCCCAT 1179
Qy 446 GlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGlyAsp 465
Db 1180 CCTACAAGTGAAGCACTTTAAGCATGAGACACAGTACGCTTATGCAA---GGCAAC 1236
Qy 466 GlyValSerSerThrAspIleGlyGlyTyrGlnLeuAlaTyrLysTrpSerGluLys 485

```


Db 1237 GCGAAAGTAGCATGGGAATGGTGGTGGTGGATATGGATATAGATACGAAACGAT 1296
 QY 486 GluGlyGluAsnGlyArgGlyGlyGlyPheLysArgValTyrLeuHisGlnGluGly 505
 Db 1297 GAA-----TACAGAGGTATTATCTTAAGAGAGATGA 1329
 QY 506 ValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGlu 525
 Db 1330 GCTGAA--TCTCGCGGTGGCTCGATCTCTATTCCCGGAGGTCCGGAT-----GGT 1380
 QY 526 GlySerGluPheValHisAlaAlaValSerGlnSerAlaLeuPheSerLysGly 545
 Db 1381 GGAGGAGCTACATTCAGCTTCCTCTTGAACGCTGTTAAAGCGGCTTGGTGTGGTCTTAATCA 1440
 QY 546 LeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLys 565
 Db 1441 -----GTTTCATGGATCCGCCATGGTTCGCCCGGAGAAAATTGCTGCTCT 1485
 QY 566 GlySerArgTyrLysAspLeuPheGluProGlyValArgArgAlaLeuValGlyVal 585
 Db 1486 GGACCACTCTGCTCTGCTCTTCTTGAACGCTGTTAAAGCGGCTTGGTGTGGTGTGTC 1545
 QY 586 GlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGln 605
 Db 1546 GGCATTCAAATACTGCAGAGTTTTCAGGTATCAATGGAGTCTCTACTACACTCCTCAG 1605
 QY 606 IleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSer 625
 Db 1606 ATTCTCGAAGCGGCTGGCTAGATATTCTTTCGAGCCTCGGACTAAGTTCCATCTCT 1665
 QY 626 AlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAla 645
 Db 1666 GCGTCAATTCCTCATACGCGTTTAAACATTTACTACTCTCTGGACAAATCCAGTCTCTC 1725
 QY 646 MetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeu 665
 Db 1726 ATGAGACTCATGGATGATCCGAGAGAGGTCAATTTACTTCTCTGGACAAATCCAGTCTCTC 1785
 QY 666 IleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHis 685
 Db 1786 ATTGCTCACTTGTCTCTCTCTGTCGAGAGTCAATCCACATCAGCAAGTCTGTAAC 1845
 QY 686 AlaLeuLeuSerThrValSerValIleValTyrPheCysPheValMetGlyPheGly 705
 Db 1846 GCAGCACTCTCCACAGGTTCTGTCGTCCTACTTCTGCTTCTTCTGATGGGTACGGT 1905
 QY 706 ProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIle 725
 Db 1906 CCCATTCCAAACATCTCTGTTCTGAAATCTTCCCAACAGAGTCCGTGGTCTCTGCATC 1965
 QY 726 AlaIleCysAlaPheThrPheThrPheIleGlyAspIleIleValThrTyrSerLeuProVal 745
 Db 1966 GCCATATGTCATGTCCTTTGGATTTGGAGACATTATTGTACAGTACTCACTTCCCGTT 2025
 QY 746 MetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIle 765
 Db 2026 CTCCTCAGCTCGATCGAGTGTGTTGTTTTCAGCATTTTACGCTCGGTTTGCATTATC 2085
 QY 766 SerPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIle 785
 Db 2086 TCATGGATCTCGTTTACATGAAGTCCCGGAGACTAAGGCATGCGCTTTGGAGTTATC 2145
 QY 786 ThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLysAla 800
 Db 2146 ACAGACTACTTGGCTTTGGAGCT---CAAGCTCAAGCTTCTGCT 2187

Search completed: October 14, 2004, 00:03:19
 Job time : 1229.17 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 13, 2004, 12:24:34 ; Search time 18.9209 Seconds
(without alignments)
2138.481 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSLAVQTFPTPLDRR.....PLEVITEFFAVGAQAAAKA 800

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 49464 seqs, 25288758 residues

Total number of hits satisfying chosen parameters: 98928

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool_h/US10051909/runat_13102004_123338_19982/app_query.fasta_1.1678
-DB=Pending Patents NA New -QWMT=fastap -SUFFIX=rrnp -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -WAXLEN=200000000
-USER=US10051909 @CGN_1_1_271 @runat_13102004_123338_19982 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents NA New:
1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	2117	52.0	94349	US-10-746-294A-96	Sequence 96, Appl
2	354	8.7	2571	US-60-613-292-1068	Sequence 1068, Ap
3	352.5	8.7	2573	US-60-613-292-899	Sequence 899, App
4	187.5	4.6	582	US-10-948-737-10147	Sequence 10147, A
5	181	4.4	1910	US-60-613-292-497	Sequence 497, App
6	166	4.1	90149	US-10-746-294A-40	Sequence 40, Appl
7	158	3.9	1882	US-60-613-292-1274	Sequence 1274, Ap
8	153	3.8	140064	US-10-746-294A-47	Sequence 47, Appl
9	142.5	3.5	1629	US-60-613-194-105	Sequence 105, App
10	141.5	3.5	411	US-60-613-194-105	Sequence 5817, Ap
11	126	3.1	9888	US-10-948-737-399	Sequence 399, App

12	123.5	3.0	473	6	US-10-950-009-135	Sequence 135, App
13	119	2.9	491	6	US-10-950-009-1142	Sequence 1142, Ap
C 14	118	2.9	2016	6	US-10-632-694A-4	Sequence 4, Appl
15	118	2.9	2016	6	US-10-632-694A-5	Sequence 5, Appl
C 16	116.5	2.9	31248	6	US-10-485-710-1	Sequence 1, Appl
C 17	116.5	2.9	35359	6	US-10-485-710-2	Sequence 2, Appl
18	112.5	2.8	237	6	US-10-950-009-605	Sequence 605, App
C 19	112	2.7	337	6	US-10-950-009-1043	Sequence 1043, Ap
C 20	112	2.7	1668	6	US-10-411-910B-226	Sequence 226, App
C 21	112	2.7	3621	6	US-10-411-910B-213	Sequence 213, App
C 22	111.5	2.7	91924	6	US-10-746-294A-48	Sequence 48, Appl
C 23	110.5	2.7	1805	1	PCT-US04-31524-92	Sequence 92, Appl
C 24	110	2.7	77040	7	US-60-612-215-1	Sequence 1, Appl
C 25	108	2.7	77040	7	US-60-612-215-1	Sequence 47, Appl
C 26	107	2.6	140064	6	US-10-746-294A-47	Sequence 91, App
C 27	106	2.6	10120	7	US-60-612-215-3	Sequence 3, Appl
C 28	104.5	2.6	5022	7	US-60-612-215-3	Sequence 244, App
C 29	104	2.6	2535	6	US-10-411-910B-244	Sequence 24, Appl
C 30	103.5	2.5	8233	6	US-10-474-148-29	Sequence 29, Appl
31	103	2.5	1129	6	US-10-950-009-370	Sequence 370, App
32	101.5	2.5	487	6	US-10-948-737-6292	Sequence 6292, Ap
C 33	101	2.5	1745	6	US-10-411-910B-223	Sequence 223, App
34	101	2.5	2627	7	US-60-613-292-514	Sequence 514, App
35	100	2.5	934	6	US-10-950-009-125	Sequence 125, App
36	100	2.5	4224	1	PCT-US04-14421-52	Sequence 52, Appl
C 37	99.5	2.4	84233	6	US-10-746-294A-105	Sequence 105, App
C 38	98.5	2.4	533	7	US-10-950-009-371	Sequence 371, App
39	98.5	2.4	1719	6	US-60-612-215-22	Sequence 22, Appl
C 40	98.5	2.4	3091	6	US-10-488-197-19	Sequence 19, Appl
C 41	97.5	2.4	1563	6	US-10-411-910B-195	Sequence 195, App
C 42	97.5	2.4	4403	1	PCT-US04-14421-76	Sequence 76, Appl
43	97.5	2.4	5034	7	US-60-612-215-4	Sequence 4, Appl
C 44	97	2.4	1911	6	US-10-411-910B-225	Sequence 225, App
C 45	96.5	2.4	2100	6	US-10-509-131-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-10-746-294A-96/c
; Sequence 96, Application US/10746294A
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Rounsley, Steven D.
; TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
; FILE REFERENCE: 38-10(15493)E
; CURRENT APPLICATION NUMBER: US/10/746,294A
; CURRENT FILING DATE: 2003-12-29
; PRIOR APPLICATION NUMBER: US 09/692,412
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 09/534,859
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 09/803,736
; PRIOR FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 124
; SEQ ID NO 96
; LENGTH: 94349
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-746-294A-96

Alignment Scores:
Pred. No.: 7.66e-187 Length: 94349
Score: 2117.00 Matches: 463
Percent Similarity: 66.02% Conservative: 114
Best Local Similarity: 52.97% Mismatches: 147
Query Match: 51.95% Indels: 153
DB: 6 Gaps: 17

US-10-051-909-32 (1-800) x US-10-746-294A-96 (1-94349)

QY 51 IleLeuGluAspIysMetSergIyAlaValLeuAlaIleValAlaSerIleGlyAsn 70

Db 53298 GTAAAGAGAAAAAGATGAGGAGTGTAGTCTGTTGCTTTAGCGGCTGGATAGGAAT 53239
Qy 71 LeuLeuGlnGlyTrpAspAsnAlaThrIleAlaala----- 82
Db 53238 ATGTTGAGGCTGGCAATGACCAATGTCAGG-TATCTCTTTCTATATTTCTCTT 53180
Qy 82 ----- 82
Db 53179 TCTTTTCTTTCATACAGGAATCTCACTATGATTTGATTTGATTTCTGGCGCA 53120
Qy 83 ---AlaValLeuTyrlleLysLysGlyPheGlnLeuGlnAsnGluProThrValGluGly 101
Db 53119 GGAGCTGTAATTTACATTAAGAAAGATTTTCATTTGGAGAAAGAACCAAGATAGAAGGA 53060
Qy 102 LeuIleValSerMetSerLeuIleGlyAlaThrIleValThrThrPheSerGlyProLeu 121
Db 53059 CTGATCGTGGCAATGCTCTCATTTGGAGCCATTTGATCACAATTTCTCTGGTCTGTC 53000
Qy 122 SerAspSerIleGlyArgProMetLeuIleLeuSerSerIleLeuTyrllePheSer 141
Db 52999 TCTGACAAATAGGAAGCGCATCAATGCTTATATCTCTCTCTCTCTCTATTTCTGAGT 52940
Qy 142 GlyLeuIleMetLeuTrpSerProAsnValTyrlleValLeuLeuAlaAargPheValAsp 161
Db 52939 AGCATGTTATGTTTGGTCTCCCATGTCATGCTTCTCTTTTCAGAGCTTCTTGAT 52880
Qy 162 GlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrlleSerGluIleAlaPro 181
Db 52879 GGTTTTGGGATCGTTAGTGTCACTCTCTGTCCTCAATCTACATCTCTGAGACCGCACCT 52820
Qy 182 SerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMetPhe 201
Db 52819 TCTGAGATCAGAGGATTAACAATTTCCCGCAGTTTGTGGATCCGGTGGGATGTTT 52760
Qy 202 LeuSerTyrlleCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMet 221
Db 52759 TTCTGTATGTTCTGTGTTTGGGATGTCGCTCAAGAAATCAACAGCTGGAGGCTGATG 52700
Qy 222 LeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrllePro 241
Db 52699 CTGGTGTGTTGTCATTCGTCATTCGTCATTTGTTGATTCGGCTTCTCTTCTGCT 52640
Qy 242 GluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGln 261
Db 52639 GAATCTCCAAGTGGCTTGTGAGAAAGCCGTGATGATGAAGCTAGACAGTTCTGCGAG 52580
Qy 262 LysLeuArgGlyLysAspValSer----- 270
Db 52579 AGACTCGTGGCAGAGAGATGTTTCAGGTTTGTCTGTGATGAACACTAAGACCTCTAGAT 52520
Qy 271 -----GlyGluLeuSerLeuLeuLe 277
Db 52519 TATCGATCTTTGAATTATTAACCTCGGGGTATGATTTAGCGAGCTGCTCTCTGCT 52460
Qy 277 uGluGlyLeuGluValGlyGlyAspThrSerIleGluGlyTyrlleIleGlyProAlaThr 297
Db 52459 TGAAGGCTTGGGGTGGGAAAGACACGTCGATAGAAATATGATGATTGACACAGCAA 52400
Qy 297 rGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrllePyr 317
Db 52399 CGAGGAAACAGAGGTGGAAATGAATCTGCCGAGGAAAGATCAGATAAACTATATGTTCC 52340
Qy 317 oGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySerVa 337
Db 52339 AGAGGATGCACAGTCATGATGCTTAAGCCAGTGAAGGACAG-----AGTTC 52292
Qy 337 lleuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspPr 357
Db 52291 TCTCGATTTGGCTTCCGACAGGAGGATGTTACCGCGTGGCGGATCCCTCATGGACCC 52232
Qy 357 oIleValThrLeuPheGlySerValHisGluAsnMetPro-----GlnAlaGlyGlySe 375

Db 52231 ACTTGTCACTCTCTTTGGTAGCATTTTCATGAGATCTCCCTTCCGAGAACATGAACGCATC 52172
Qy 375 rMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHisAl 395
Db 52171 ATCCGCGAGCATGCTCTTCCCAATATGGAAGTATATCTGGGAATGATGGAAGG----- 52117
Qy 395 alyAsnGlnGlnTrpAspGluAsnLeuHisArgAspAspGluGluTyrlleAlaSerAs 415
Db 52116 -CAGAGTCCCACTGGGATCCAGAG-----AGAAACAATGAACAT----- 52078
Qy 415 pGlyAlaGlyGlyAspTyrlleAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThr 435
Db 52077 -----AGTTCTGACCAAGATGAATACTTAACAGATCTCTGCTTTCTCCGCAACAC 52025
Qy 435 rGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMetAr 455
Db 52024 T-----GAGCGGATGAC-----TACCACCGCGACCGTGGTACCATGCTAGCGG 51977
Qy 455 gArgGlnThrLeu---LeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGly 474
Db 51976 ACAGAGCAGCTGTTGTTATGCAAAACGTGGGTGAGACAGACGCGCTACAAGCATAGCGG 51917
Qy 474 YGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLys---G 493
Db 51916 TGGATGGCAATGTCGTTGGAGATATAATGACAAGTTGGTGGCAGATGGTAAGAGAGTCAA 51857
Qy 493 uGlyGlyPheLysArgValTyrlleHisGlnGlu-----GlyVa 506
Db 51856 TGGAGGTTACAGAGAATGATATTCACGAGGAAACCGCCAAACAAACACCAACACAT 51797
Qy 506 lProGlySerArgArgGlySerIleValSerLeu---ProGlyGlyGlyAspValPheG 525
Db 51796 TCTTTTTCGCGAGCTGGATCACTCTCTCTCCACCCAGCGGTGATGGTCAATGATCA 51737
Qy 525 uGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysG 545
Db 51736 GGTGAATGGATACGTTGAGCTGCTGACCTGTGAGCAAGCTTCGATGATGCCAGAGG 51677
Qy 545 YLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAla 565
Db 51676 TAAAGCGCAGACCGCAATGTTG-----CCAAAGGAGGTT---AAGGA 51638
Qy 565 sGlySerArgTrpIleAspLeuPheGluProGlyValArgAlaLeuValGlyVa 585
Db 51637 TGGTCCAGGCTGGAGGAGCTGAAAGAACCCAGGGGTTAAGAGAGCTTTGATGGTTGAGT 51578
Qy 585 lGlyIleGlnIleLeuGln----- 591
Db 51577 TGGGCTTCAGATACTGCA-ACAGGTAATGTTTTTTTCCAGATCTTTGCTGTAGTAAAG 51519
Qy 592 -----GlnPheAlaGlyrl 596
Db 51518 AGACAAACAAATAAGAACTAAATAACGGAAGAAAGCATGTTTCAGTTCAGGAAT 51459
Qy 596 eAsnGlyValLeuTyrlleThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLe 616
Db 51458 AAATGGAGTGTATATATACACCTCAATATTTGGAAGAAACAGGTTGTCTCAAGTCTTT 51399
Qy 616 uSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrIle 636
Db 51398 GACAAACCTTGGAAATAGTGAGAGTCTGCTGCTCTTCATAAGCGCTTAACACACT 51339
Qy 636 uLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArph 656
Db 51338 CTGTGATGCTTCCCTGATTTCTGCTCCATGAGTTAATGGATGTAACTGGCAGAGGCG 51279
Qy 656 e----- 656
Db 51278 -AGTTTCTCTGCTTCCCTTCTCTTTTATAGTGAATGGTATTAATTTGAAAGATPA 51220
Qy 657 -----LeuLeuLeuGlyThrIleProIleLeuIleAl 667
Db 51219 ACATACCAAGAGCGTGAATGCGAGGCTCTGATGCTTTCGACTATCCCATCTTAATACT 51160

QY 667 aSerLeuValIleLeuValValSerAsnLeuIleLeuLeuGlyThrLeuAlaHisAlaLe 687
DB 51159 GTCCCTGGTAACACTCGTGAAGAGCTTAGTGAATCTTGAGGTTCAATAAAGCGGT 51100
QY 687 uLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProI 707
DB 51099 GATATCGACGCAAGTGTACGGTGTAACCTAGCTGTTTCGTGATGGGTTTGGGGCAAT 51040
QY 707 eProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaI 727
DB 51039 TCCAAACATCTCTGCTCAGAGATATTCCTACTTCTGCGCGGCTCTGATCACCAT 50980
QY 727 eCysAlaPheThrPheTrpIleGlyAspIleValThrTyrSerLeuProValMetLe 747
DB 50979 ATGTGCGCTCACCCTTCGTGATCTGTGACATAATCGTCACTTACACCTCCAGTCATGT 50920
QY 747 uAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPh 767
DB 50919 CAAATCCATTGGCAGAGAGTCTTGGCAATTTATGCAATCGTCTGCTGCTGCGGTG 50860
QY 767 eValPheValPheLeuIysValProGluThrLysGlyMetProLeuGluValIleThrG 787
DB 50859 GGTGTTTGTGACTCTGAGTACAGACAGACAAAGGGAATGCCCTTGAAGTTATCTCTGA 50800
QY 787 uPhePheAlaValIleAlaLysGlnAlaAlaLysAla 800
DB 50799 GTTCTCTCCGTCGGTGCAAAACAGCAAGACGCTGCAGCT 50760

RESULT 2

US-60-613-292-1068
; Sequence 1068, Application US/60613292
; GENERAL INFORMATION:
; APPLICANT: Gene Logic, Inc
; APPLICANT: HIGGS, Brandon
; APPLICANT: ELASHOFF, Michael
; APPLICANT: MENDRICK, Donna L.
; APPLICANT: PORTER, Mark W.
; APPLICANT: CASTLE, Arthur L.
; APPLICANT: JOHNSON, Kory R.
; TITLE OF INVENTION: MOLECULAR TOXICITY MODELS FROM ISOLATED HEPATOCYTES
; FILE REFERENCE: 044921-5135-PR
; CURRENT APPLICATION NUMBER: US/60/613,292
; CURRENT FILING DATE: 2004-09-28
; NUMBER OF SEQ ID NOS: 1300
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1068
; LENGTH: 2571
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION:
US-60-613-292-1068

Alignment Scores:
Pred. No.: 2,76e-24 Length: 2571
Score: 354.00 Matches: 155
Percent Similarity: 33.41% Conservative: 119
Best Local Similarity: 18.90% Mismatches: 205
Query Match: 8.69% Indels: 341
DB: 21 Gaps: 21

US-10-051-909-32 (1-800) x US-60-613-292-1068 (1-2571)

QY 4 GlySerTrpLeuAlaValGlnThrProPheThrProAspLeuAspArgGluArgLeu 23
DB 133 GGGAGAGCATAGTCACAGACGTCCTTCTCC----- 165
QY 24 LeuProSerValValLeuAlaLeuProGlyProLeuProProAlaSer-CysSerSerG 43
DB 166 -----GTTTCACAGCCCGCACAGCTTGAGCTTCGAGCGCGGCCAT 210
QY 43 nGluProValThrSerAspIleLeuGluAspIysMetSerGlyAlaValLeuValAl 63

DB 211 GGAGCCCGCAGCAAG-----AAGGTGACGGCGCGCTTATGTTGGC 252
QY 63 aile--ValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIleAlaAl 82
DB 253 CGTGGAGGGCAGTGTCTGGATCCCTGCGCTTAACACCGGTGTCTCAACGC 312
QY 82 aAlaValLeuTyrIleLysLysGluPhe---GlnLeuGlnAsnGlu----- 96
DB 313 CCCCAGAAAGGTAATTGAGGAGTTCTACAATCAACATGGAACCCCGCTATGGAGATC 372
QY 97 -----ProThrValGluGlyLeuIleValSerMetSerLeuIleG 110
DB 373 CATCCATCCACACATCACCACACTCTGGTCTCTCTCCGTGGCCATCTCTCTGTCGG 432
QY 110 yAlaThrIleValThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMe 130
DB 433 GGGCATGATTGTTCCCTCTCTCTGGGCTCTTTGTTAATCGCTTGGCAGCGGAATC 492
QY 130 tLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSer--- 148
DB 493 CATGTGATGAAACCTGTTGGCTTTGTCTGCCGTGCTTATGGTTTCTCAAACT 552
QY 149 ---ProAsnValTyrValLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLe 167
DB 553 GGGCAAGTCCCTTGAGATGCTGATCTCTGGCGCTTCATCATTCGAGTGTACTGTGGCT 612
QY 167 uAlaValThrLeuValProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLe 187
DB 613 GACCACCGGCTTTGTGCCCATGTATGTGGGAGGTGTCACCCACAGCTCTCTGGGAGC 672
QY 187 uLeuAsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeuSerTyrCysMetVa 207
DB 673 CCGTGGCACCTTGACACAGCTGGGCATCGTCTGGGATCTTATGGC-----CAGGT 726
QY 207 lPheGlyMet---SerLeuSerProSerProAsp---TrpArgIleMetLeuGlyValLe 225
DB 727 GTTCGGCTTAGACTCCATCATGGGCAATGCAGACTTGTGGCTCTACTGCTCAGTGTCA 786
QY 225 uAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuProGluSerProAr 245
DB 787 CTTTCATCCAGCCCTGCTACAGTGT---ATCTGTGGCTTCTCCCTCGAGAGCCCGC 843
QY 245 gTrp---LeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuAr 264
DB 844 CTTCTGCTCATCAATCGTAACGAGAGAACCGGGCCAAAGAGTGTGCTGCTGAAAGCTTCG 903
QY 264 gGlyLysAspAspValSerGlyLeuSerLeuLeuGluGlyLeuGluValGlyG 284
DB 904 AGGACAGCCGATGTGACCCGAGACCTGACGAGATGAAAGAA----- 946
QY 284 yAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAlaAspLeuVa 304
DB 946 ----- 946
QY 304 lThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIl 324
DB 946 ----- 946
QY 324 eAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHi 344
DB 946 ----- 946
QY 344 sGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySe 364
DB 946 ----- 946
QY 364 rValHisGluAsnMetProGlnAlaGlySerMetArgSerThrLeuPheProAsnPh 384
DB 946 ----- 946
QY 384 eGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAs 404

Db 946 ----- 946
 Qy 404 nLeuHisArgAspGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAs 424
 Db 946 ----- 946
 Qy 424 nLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHi 444
 Db 947 ----- 947
 Qy 444 sHisGlyHisArgGlySerAlaLeuSerMetArgGlnThrLeuLeuGlyGluGlyG 464
 Db 967 GGAG-----AGAGGTCACCATCTTG----- 968
 Qy 464 yAspGlyValSerThrAspIleGlyGlyTyrGlnLeuAlaTrpLysTrpSerG 484
 Db 988 ----- 988
 Qy 484 uLysGluGlyGluasnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnG 504
 Db 988 ----- 988
 Qy 504 uGlyValProGlySerArgGlySerIleValSerLeuProGlyGlyGlyAspValPh 524
 Db 988 ----- 988
 Qy 524 eGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLy 544
 Db 988 ----- 988
 Qy 544 sGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAl 564
 Db 988 ----- 988
 Qy 564 aLysGlySerArgTrpLysAspLeuPhe---GluProGlyValArgArgAlaLeuValVa 583
 Db 989 ----- 989
 Qy 583 lGlyValGlyIleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyrThr 603
 Db 1030 CGCGTGGTCTCGAGTGTCCAGAGCTGTCCGGCATCAATGCTGTCTTACTACTTC 1089
 Qy 603 rProGlnIleLeuGluAlaGlyVal---AlaValIleLeuSerLysPheGlyLeuSe 622
 Db 1090 AACGAGCATCTTCGAGAAGCAGGTGTGCAGCAGCCTGTGTATGCCACCATCGGCTCGG 1149
 Qy 622 rSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysAl 642
 Db 1150 TATCGTCAACACGCGCTTCACTGTGTGTGCTG----- 1183
 Qy 642 eGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeu---LeuLeuGlyTh 661
 Db 1184 -----TTCGTGTGTGAGCAGCTGGCGTGTGCAGCCCTGCATCTATTGGTCT 1230
 Qy 661 rIleProIleLeu---IleAlaSerLeuValIleLeuValValSerAsnLeuIleAspIe 680
 Db 1231 GGCTGCATGGCGGTGTGTGTCTATGACCATCGCCCTGGCCCTGCTGCAGCAGCT 1290
 Qy 680 uGlyThrLeuAlaHisAlaLeuSerThrValSerValIleValTyrPheCysCysPh 700
 Db 1291 GCCTGTGATGTCCTAT-----CTGAGTATCGTGGCCATCTTGGCTTGTGGCTTCTT 1344
 Qy 700 eValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgVa 720
 Db 1345 TGAAGTAGCCCTGTGTCCTATTCCATGTTCTATGTGGCGGAGCTGTTCAGCCAGGCGCC 1404
 Qy 720 lArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValTh 740
 Db 1405 CCGACCTGT 1464
 Qy 740 rTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAl 760
 Db 1465 CATGTGCTTCCCAATATGTGGAGCAACTGTGTGGC---CCCTACGTCTTTCATCATCTTCC 1521

Qy 760 aValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMe 780
 Db 1522 GGTGCTGTGGTACTTCTTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 1581
 Qy 780 tProLeuGluValIleThrGluPheAlaValAlaGlyAlaLysGlnAlaAlaLys 799
 Db 1582 GACCTTCATGATGATGCTTCCGGCTTCCGGCAGGGGGGTGCCACGACGCGCAAG 1639

RESULT 3

US-60-613-292-899
 ; Sequence 899, Application US/60613292
 ; GENERAL INFORMATION:
 ; APPLICANT: Gene Logic, Inc
 ; APPLICANT: HIGGS, Brandon
 ; APPLICANT: ELASHOFF, Michael
 ; APPLICANT: MENDRICK, Donna L.
 ; APPLICANT: PORTER, Mark W.
 ; APPLICANT: CASTLE, Arthur L.
 ; APPLICANT: JOHNSON, Kory R.
 ; TITLE OF INVENTION: MOLECULAR TOXICITY MODELS FROM ISOLATED HEPATOCYTES
 ; FILE REFERENCE: 044921-5135-PR
 ; CURRENT APPLICATION NUMBER: US/60/613,292
 ; CURRENT FILING DATE: 2004-09-28
 ; NUMBER OF SEQ ID NOS: 1300
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 899
 ; LENGTH: 2573
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; OTHER INFORMATION:
 US-60-613-292-899

Alignment Scores:

Pred. No.: 3.82e-24 Length: 2573
 Score: 352.50 Matches: 157
 Percent Similarity: 31.68% Conservative: 105
 Best Local Similarity: 18.98% Mismatches: 212
 Query Match: 8.65% Indels: 353
 DB: 7 Gaps: 18

US-10-051-909-32 (1-800) x US-60-613-292-899 (1-2573)

Qy 1 IleArgSerGlySerTrpLeuAlaValGlnThrProPheThrProAspLeuAspArgArg 20
 Db 65 ATTGGAGACAGGAGCTGGATTGAGACACAACTCCGACACACATGTTCAGACACAGA 124
 Qy 21 GluArgLeuProSerValValLeuAlaLeuProGlyProLeuProProAlaSerCys 40
 Db 125 -----TCACCGAAGCTTGGCTTTCACGTG 148
 Qy 41 SerSerGlnGlu-----ProValThrSerAsp---AspIle----- 51
 Db 149 TCCTTCACTGCAGTGTGGGTTCCTTCCAGTTCGGATATGACATCGGTGTGATCAATGCAC 208
 Qy 52 -----LeuGluAspLys- 55
 Db 209 CTCAAGAGGTAAATAATATCCATTATCGACATGTTTGGGTGTCTCTGTGATGACCGAA 268
 Qy 56 -----MetSerGlyAlaValLeuValAlaIleValAlaSerIleG 69
 Db 269 GAGCTACCATTAAGTATGACATCAATGGCAGACAGACCCCACTCATAGTCACACAGCAC 328
 Qy 69 lYasnLeuLeuGlnGlyTrpAspAsnAlaThrIleAlaAlaValLeuTyrIleLysL 89
 Db 329 ATACGACACACAGACGCTCGGAAGAAGAGACTGCAAGATCTGCTCATATAGTC----- 381
 Qy 89 ysgLupheGlnLeuGlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeu 109
 Db 382 -----ACTATGCTCTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 418
 Qy 109 leGlyAlaThrIleValThrPheSerGlyProLeuSerAspSerIleGlyArgArgP 129

; APPLICANT: HANSEN, RHONDA
 ; APPLICANT: KAUFMANN, JOERG
 ; APPLICANT: KENNEDY, GIULIA C.
 ; APPLICANT: LAMSON, GEORGE
 ; APPLICANT: MOLER, EDWARD J.
 ; APPLICANT: RANDAZZO, FILIPPO
 ; APPLICANT: REINHARD, CHRISTOPH
 ; APPLICANT: SUDDUTH-KLINGER, JULIE
 ; TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
 ; TITLE OF INVENTION: IN CANCEROUS CELLS III
 ; FILE REFERENCE: 2300-21987
 ; CURRENT APPLICATION NUMBER: US/10/948,737
 ; CURRENT FILING DATE: 2004-09-22
 ; PRIOR APPLICATION NUMBER: 10/616,900
 ; PRIOR FILING DATE: 2003-07-09
 ; PRIOR APPLICATION NUMBER: 09/872,850
 ; PRIOR FILING DATE: 2001-06-01
 ; PRIOR APPLICATION NUMBER: 60/208,871
 ; PRIOR FILING DATE: 2000-06-02
 ; PRIOR APPLICATION NUMBER: 10/081,519
 ; PRIOR FILING DATE: 2002-02-21
 ; PRIOR APPLICATION NUMBER: 60/270,959
 ; PRIOR FILING DATE: 2001-02-21
 ; PRIOR APPLICATION NUMBER: 10/310,673
 ; PRIOR FILING DATE: 2003-01-08
 ; PRIOR APPLICATION NUMBER: 60/336,613
 ; PRIOR FILING DATE: 2001-12-04
 ; PRIOR APPLICATION NUMBER: US03/00657
 ; PRIOR FILING DATE: 2003-01-08
 ; PRIOR APPLICATION NUMBER: 60/345,637
 ; PRIOR FILING DATE: 2002-01-08
 ; PRIOR APPLICATION NUMBER: 10/081,124
 ; PRIOR FILING DATE: 2002-02-21
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 13996
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 10147
 ; LENGTH: 582
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-948-737-10147

Alignment Scores:
 Pred. No.: 1,06e-09 Length: 582
 Score: 187.50 Matches: 38
 Percent Similarity: 55.56% Conservative: 22
 Best Local Similarity: 35.19% Mismatches: 45
 Query Match: 4.60% Indels: 3
 DB: 6 Gaps: 1

US-10-051-909-32 (1-800) x US-10-948-737-10147 (1-582)

QY 687 LeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyPro 706
 Db 16 CTGCTGCCACCC-----ATGCTCTTCATCANGGCTACGCCGTGGGCTGCCGCC 66
 QY 707 IleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAla 726
 Db 67 ATCACCTGGCTGTCATGCTGAGGCTCCGCCCTGCGCGCGGCTGCGCTCAGG 126
 QY 727 IleCysAlaPheThrPheThrIleGlyAspIleIleValThrTyrSerLeuProValMet 746
 Db 127 CTCTGGCTGTCGCCAGCTGGCTCACCGCTTGTCTCCACCAAGTCTTCTCGCCAGTG 186
 QY 747 LeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSer 766
 Db 187 GTGAGCACCTTGGCTCCAGGTGCCCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTT 246
 QY 767 PheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThr 786
 Db 247 CTGGTGTTCACAGCTGCTGTGTGCCGAGACCAAGGAGCGGCTCCCTGGAGCAGATCGAG 306
 QY 787 GluPhePheAlaValGlyAlaLys 794

Db 307 TCCTTCTTCCGACGGGAGAGG 330
 RESULT 5

US-60-613-292-497
 ; Sequence 497, Application US/60613292
 ; GENERAL INFORMATION:
 ; APPLICANT: Gene Logic, Inc
 ; APPLICANT: HIGGS, Brandon
 ; APPLICANT: ELASHOFF, Michael
 ; APPLICANT: MENDRICK, Donna L.
 ; APPLICANT: PORTER, Mark W.
 ; APPLICANT: CASTLE, Arthur L.
 ; APPLICANT: JOHNSON, Kory R.
 ; TITLE OF INVENTION: MOLECULAR TOXICITY MODELS FROM ISOLATED HEPATOCYTES
 ; FILE REFERENCE: 044921-5135-PR
 ; CURRENT APPLICATION NUMBER: US/60/613,292
 ; CURRENT FILING DATE: 2004-09-28
 ; NUMBER OF SEQ ID NOS: 1300
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 497
 ; LENGTH: 1910
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; OTHER INFORMATION:
 ; US-60-613-292-497

Alignment Scores:
 Pred. No.: 2,69e-08 Length: 1910
 Score: 181.00 Matches: 75
 Percent Similarity: 44.97% Conservative: 59
 Best Local Similarity: 25.17% Mismatches: 98
 Query Match: 4.44% Indels: 66
 DB: 7 Gaps: 15

US-10-051-909-32 (1-800) x US-60-613-292-497 (1-1910)

QY 23 LeuLeuProSerValValLeuAlaLeuProGly-----ProLeuProPro 37
 Db 179 CTCCTGCCCTCTTCATGGCGCGTGTGCTGCCACCATTTGCTGCCCTGCCCTGCCCC 238
 QY 38 AlaSerCysSerSerGlnGlu-----ProValThrSerAsp 50
 Db 239 GCCAACCTCAGTCACAGGACTTATGGCTGGAAGCCCTACTACCCGGGAGACTGAC--- 295
 QY 51 IleLeuGluAspLysMetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsn 70
 Db 296 -----GGCAGCTTTAGCTCTGCTCCGATTTGCTTATCCCAAGACTGTCCCAAT 346
 QY 71 LeuLeu----- 72
 Db 347 GTCACTTTGGGACAGAGGTGTCCAACTCTGGGAGCCTGAGGCTGAGCCCTCACGGTG 406
 QY 73 -----GlnGlyTrpAsp-----AsnAlaThrIleAlaAla 83
 Db 407 CCCTGCTCTCAGGCGTGGAGTACGACCGCTCAGAAATTCCTCCACCATTTGCAACTGAG 466
 QY 84 ValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeuIle 103
 Db 467 TGGGATCTTGTGTGTGTCAGAGAGAGACTG---AACAAATTTACGTCACCTCTTCTTC 523
 QY 104 ValSerMetSerLeuIleGlyAlaThrIleValThrPheSerGlyProLeuSerAsp 123
 Db 524 ATTTGGTGTG---CTGGTGGGAGCGGTGTAT-----GGATATTGTCTGAC 568
 QY 124 SerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeu 143
 Db 569 AGGTTTGGCAGGCGCGGCTTCTGCTGGTGGCTACGTGAGCTCCCTGGTGTGCTGCTG 628
 QY 144 IleMetLeuTrpSerProAsnValTyrValLeuLeuAlaArgPheValAspGlyPhe 163
 Db 629 ATGTCTGCACCTCCATCAACTACATCATGTTCTGTAGTACCCTGACTACCTACCGGCTCA 688

Db	38775	CAC	TGCACATCATTTCCGGCTCTGGCAGTGTAGTGGTGTCTTGGTAAACATCTATAAACCTA	38834
		:::		
Qy	679	sP	euGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValle	694
		::		
Db	38835	ATTTGGTTTCAATATATATGTTTGGTCTTGGTGT	38894	
Qy	695	-----Val	tyrPheCysCysPheValMetGlyPheGlyProIleProAenIleLeuC	712
		----	-----	-----
Db	38895	ATTTCAGGTCTACATCGGTCGTTCTCTATTGGATGGGTGCAATCCCATGGGTATAA	38954	
Qy	712	ys	alaGlu	714
		:::	-----	-----
Db	38955	TGTC	TGAGGTAATTTCAAAACAATATATAATGTCATATTTTCATTATTTTTCCTTAATCT	39014
Qy	715	-----	-----	-----
		-----	-----	-----
Db	39015	TGTC	ATTACATTAATTTACTCTTTTATCTTCAGATATTTCCATAAATTTTGAAG	39074
Qy	722	ly	LeuCysIleAlaIleCysAlaPheThrPheTyrIleGlyAspIleleValThrTyrS	742
		:::		
Db	39075	GAACTGCGGAGGACTTGTCAC	TGTAGTAAATGGCTAAGTTCATGGCTTGCTCATTTCA	39134
Qy	742	er	-----	-----
		::	-----	-----
Db	39135	CTTTC	CAACTTCTCATGATTTGGAGCCCTCATGTATATATACATTTTTTTTGTATTATTT	39194
Qy	743	-----	-----	-----
		-----	-----	-----
Db	39195	ATAT	ACATAACTACTAGTGGAAACATCAGTATTGATCGCATTTCTTTGTTTGTAGGTACA	39254
Qy	756	Ph	eSerIleTyAlaValValCysIleleSerPheValPheValPheLeuIysValPro	775
Db	39255	TTTTATGTGTATGGTGGGTATGCTATGGCTATTATCTTCATAGCAAACTTGTTCT	39314	
Qy	776	Gl	uThrLysGlyMetProLeuGluValIle	785
Db	39315	GAA	CTAAAGGCAGAACCTTGAAGAGATT	39344
RESULT 9				
US-60-613-194-105				
; Sequence 105, Application US/60613194				
; GENERAL INFORMATION:				
; APPLICANT: BASF Plant Science				
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING PROTEINS ASSOCIATED WITH				
; TITLE OF INVENTION: STRESS RESPONSE AND PLANT CELLS AND PLANTS WITH INCREASED				
; FILE REFERENCE: AE2004-0486				
; CURRENT APPLICATION NUMBER: US/60/613,194				
; CURRENT FILING DATE: 2004-09-24				
; NUMBER OF SEQ ID NOS: 220				
; SOFTWARE: PatentIn version 3.2				
; SEQ ID NO 105				
; LENGTH: 1629				
; TYPE: DNA				
; ORGANISM: Saccharomyces cerevisiae				
; FEATURE:				
; NAME/KEY: CDS				
; LOCATION: (1)..(1629)				
US-60-613-194-105				
Alignment Scores:				
Pred. No.:	8,44e-05	Length:	1629	
Score:	142.50	Matches:	59	
Percent Similarity:	40.80%	Conservative:	43	
Best Local Similarity:	23.60%	Mismatches:	97	
Query Match:	3.50%	Indels:	51	
DB:	7	Gaps:	8	
US-10-051-909-32 (1-800) x US-60-613-194-105 (1-1629)				
Qy	17	Leu	AsPargGluArgLeuProSerValValLeuAlaLeuProGlyProLeuPro	36
		:::		

103 TTGACTAGGACCGAGAGTGTAAAGCCAGACCGAGATAACC-----GCTCGC 150
 Qy 37 ProAlaSerCysSerSerGlnGluProValThrSerAspAlaLeuGluAspLysMet 56
 Db 151 CCTCACTACCGCTTTTCCCGCTTTTCAAGACA-----
 Qy 57 SerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAsp 76
 Db 184 -----GTTTATAGCTCAGTCGCTTTTCACTGGGTTTTTC----- 219
 Qy 77 AsnAlaThrIleAlaAlaValLeuTyr-----IleLysLysGluPhe 91
 Db 220 ---TCCACATAGCAGGTGCCATCTACTACTCCAGTTCTCGAGCGTTATAGAAAGAAATTC 276
 Qy 92 GlnLeuGlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAla 111
 Db 277 GATATTGACGAGGAA-----TTGGTGAATGTC 303
 Qy 112 ThrIleValThr-----ThrPheSerGlyProLeuSer 122
 Db 304 ACTGTTGTAGTATATTTTGTATTTCAGGGTCTTGGCCCACTTCATGGCGGTTTGC 363
 Qy 123 AspSerIleGlyArgArgProMetLeuLeuSerSerIleLeuTyrPhePheSerGly 142
 Db 364 GATTCACTGGCAGGAGCGGTGGTGTCTGCGCAATCGTCATTTATTTTGGTGCCTGC 423
 Qy 143 LeuIleMetLeuTrpSerProAsnValTyrValLeuLeuAlaArgPheValAspGly 162
 Db 424 ATCGGCTTCTGCTGTCTCAAAAGTATGCTCAGATCATTTGCTAAGATGCTACAAGCC 483
 Qy 163 PheGlyGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaPro--- 181
 Db 484 GCGCGTATTTCACCGGTGATGCGGATTAACAGCGGAATAATGCGAGATGTTACTACTAGA 543
 Qy 182 SerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMetPhe 201
 Db 544 CGCAGCGCGCGGTAGCTGTGATATGTTGCGATTTCAAGTCTAGTCTCGGTTTC 603
 Qy 202 LeuSerTyrCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMet 221
 Db 604 GGAGCC-----CTTATCGGTGCGGATATATCATCTAGATGGGATGAGAGCCATC 654
 Qy 222 LeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuPro 241
 Db 655 TTTTGGTCTTACAAATGATCGCAATTTGCTTCTAGCCTCGTTTAAATTTTGCCA 714
 Qy 242 GluSerProArgTrpLeuValSerLysGly 251
 Db 715 GAAACAAAGAGAACATATCCGGGAATGTT 744

RESULT 10

US-10-948-737-5817/c
 ; Sequence 5817, Application US/10948737
 ; GENERAL INFORMATION:
 ; APPLICANT: CHAN, VIVIEN W.
 ; APPLICANT: ESCOBEDO, JAIME
 ; APPLICANT: GARCIA, PABLO DOMINGUEZ
 ; APPLICANT: HANSEN, RHONDA
 ; APPLICANT: KAUFMANN, JOERG
 ; APPLICANT: KENNEDY, GIULIA C.
 ; APPLICANT: MOLER, EDWARD J.
 ; APPLICANT: RANDAZZO, FILIPPO
 ; APPLICANT: REINHARD, CHRISTOPH
 ; APPLICANT: SUDDUTH-KLINGER, JULIE
 ; TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
 ; FILE OF INVENTION: IN CANCEROUS CELLS 111
 ; FILE REFERENCE: 2300-21987
 ; CURRENT APPLICATION NUMBER: US/10/948,737
 ; CURRENT FILING DATE: 2004-09-22
 ; PRIOR APPLICATION NUMBER: 10/616,900
 ; PRIOR FILING DATE: 2003-07-09
 ; FILE REFERENCE: 09/872,850

; PRIOR FILING DATE: 2001-06-01
 ; PRIOR APPLICATION NUMBER: 60/208,871
 ; PRIOR FILING DATE: 2000-06-02
 ; PRIOR APPLICATION NUMBER: 10/081,519
 ; PRIOR FILING DATE: 2002-02-21
 ; PRIOR APPLICATION NUMBER: 60/270,959
 ; PRIOR FILING DATE: 2001-02-21
 ; PRIOR APPLICATION NUMBER: 10/310,673
 ; PRIOR FILING DATE: 2003-01-08
 ; PRIOR APPLICATION NUMBER: 60/336,613
 ; PRIOR FILING DATE: 2001-12-04
 ; PRIOR APPLICATION NUMBER: US03/00657
 ; PRIOR FILING DATE: 2003-01-08
 ; PRIOR APPLICATION NUMBER: 60/345,637
 ; PRIOR FILING DATE: 2002-01-08
 ; PRIOR APPLICATION NUMBER: 10/081,124
 ; PRIOR FILING DATE: 2002-02-21
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 13996
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5817
 ; LENGTH: 411
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-948-737-5817

Alignment Scores:
 Pred. No.: 1,24e-05 Length: 411
 Score: 141.50 Matches: 35
 Percent Similarity: 56.78% Conservative: 32
 Best Local Similarity: 29.66% Mismatches: 44
 Query Match: 3.47% Indels: 7
 DB: 6 Gaps: 4

US-10-051-909-32 (1-800) x US-10-948-737-5817 (1-411)

Qy 118 SerClyProLeuSerAspSerIleGlyArgProMetLeuIleLeuSerSerIleLeu 137
 Db 348 ACTGGCTGACCTGTGACCCGCTCCCGAGCGGAATTCATGATGATGAACCTGCGC 289
 Qy 138 TyrPhePheSerGlyLeuIleMetLeuTrpSer-----ProAsnValTyrValLeu 154
 Db 288 GCCTTCGTGTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 229
 Qy 155 LeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuProLeu 174
 Db 228 ATCTGCGCGCTTCATCATCGGTGTACTGCGGCTGACACAGCTTCGTCGCCCATG 169
 Qy 175 TyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPhe 194
 Db 168 TATGTGGTGAAGTGTACCCACAGCCCTTCGTGGGCTGCGGCTGCGGCTGCGGCTG 109
 Qy 195 SerGlySerGlyMetPheLeuSerTyrCysMetValPheGlyMet---SerLeuSer 213
 Db 108 GGCATCGTCGCGCATCTCATCGCC-----CAGGTGTGCGGCTGCGGCTGCGGCTG 55
 Qy 214 ProSerProAsp---TyrArgIleMetLeuGlyValLeuAlaIleProSerLeu 230
 Db 54 GGCAACAGGACCTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1

RESULT 11
 US-60-613-292-399
 ; Sequence 399, Application US/60613292
 ; GENERAL INFORMATION:
 ; APPLICANT: Gene Logic, Inc
 ; APPLICANT: HIGGS, Brandon
 ; APPLICANT: ELASHOFF, Michael
 ; APPLICANT: MENDRICK, Donna L.
 ; APPLICANT: PORTER, Mark W.
 ; APPLICANT: CASTLE, Arthur L.
 ; APPLICANT: JOHNSON, Kory R.
 ; TITLE OF INVENTION: MOLECULAR TOXICITY MODELS FROM ISOLATED HEPATOCYTES
 ; FILE REFERENCE: 044921-5135-PR


```

Qy 544 -----LysGlyLeuAlaGluProArgMetSerAspAla 555
Db 2063 AGACAGAGATGTGAGACTGATGTGAAGCCCTGGCCCTCAGCTGTATTGGTCAGCTGT 2122
Qy 555 aMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluPr 575
Db 2123 GGCCCTTATCCAGAGTCGTTCTTCAGAAA----- 2153
Qy 575 oGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeu-----G 591
Db 2154 -----CTCTACAAAGTACCTCTCAGTACCATGCAAGTACTGAGGA 2194
Qy 591 nGlnPheAlaGly-----IleAsnGlyValLeuTyrTrpProGlnIleLeuGlu 609
Db 2195 ACAGTAGTCTCTGACATCTCGACTACATCGCATCGAGCCCTCAGGTG----- 2246
Qy 609 nAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerIleLe 629
Db 2247 -CGAGGAGCTACTGCCATTCTC-----TGTTGGGACCCCT 2278
Qy 629 uIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMe 649
Db 2279 TGCTACTCATCTCTCAGCAGGTCGCTCTCCGTGTGGTGTGCTGCTGGCACCATCAG 2338
Qy 649 tAspLeuSerGlyArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLe 669
Db 2339 GGCCCTGACAGGAATACATTCTCTGTGTGGACTGCATTCTCTTACTGCAGAAAACTTT 2398
Qy 669 u-----ValIleLeuValValSerAsnLeuIleAspLeu---GlyThrLeuAlaH 685
Db 2399 GAAGGATGAATCTCTGTACTTGCAAGTGGCTGTGACAGCTGTGAGCAGCTGTCTCT 2458
Qy 685 iAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheVal 701
Db 2459 GAGTCITTTGACGAGCAGCTACAGTCACTTGGGATTACAACTGCTTATTG 2508

RESULT 12
US-10-950-009-135
; Sequence 135, Application US/10950009
; GENERAL INFORMATION:
; APPLICANT: BERKA, Randy
; APPLICANT: BASHKIROVA, Elena
; APPLICANT: REY, Michael
; TITLE OF INVENTION: Methods For Monitoring Gene Expression
; FILE REFERENCE: 10541.200-US
; CURRENT APPLICATION NUMBER: US/10/950,009
; PRIOR FILING DATE: 2004-09-24
; PRIOR APPLICATION NUMBER: 60/506,140
; NUMBER OF SEQ ID NOS: 1190
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 135
; LENGTH: 473
; TYPE: DNA
; ORGANISM: Trichoderma reesei
US-10-950-009-135
Alignment Scores:
Pred. No.: 0.000748 Length: 473
Score: 123.50 Matches: 45
Percent Similarity: 43.29% Conservative: 26
Best Local Similarity: 27.44% Mismatches: 58
Query Match: 3.03% Indels: 35
DB: 6 Gaps: 6

US-10-051-909-32 (1-800) x US-10-950-009-135 (1-473)
Qy 576 GlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPheAlaGly 595
Db 5 GGTGGCGGCGAGGTACACTTGTGGCGGTGAGTGCAGTGCAGATTTGGCAGCACTCTCGGT 64
Qy 596 IleAsnGlyValLeuTyrTrpProGlnIleLeuGluAlaGlyValAlaValIle 615

```

```

Db 65 GGTAAATGTGATGCTTTACTACTTGTCTACATCTTCAACATGGCTGGCATG----- 115
Qy 616 LeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThr 635
Db 116 -----TCCGGAATATACCGGCTCAGCTCTTCAATATCCCAATAT 154
Qy 636 LeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg 655
Db 155 GTATCTTTCTGTCACGACTGGAGGAGTCTCTTTGCTGTTGATCGCATTTGGCGCG 214
Qy 656 PheLeuLeuLeu-----GlyThrIleProIleLeuIleAlaSerLeu 669
Db 215 TGGTGTCTTATTGTGTGGGCGCATCATTTGGGCGCTCATCTTTCATCGTCCGTGCC-- 271
Qy 670 ValIleLeuValValSerAsnLeuIleAspLeu----- 680
Db 272 GTCAATGCGAGTCTACGGGCGCACCGTCGACAGTGTGGATGGAACGATATCTGAGGTGG 331
Qy 681 -----GlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPhe 697
Db 332 CAGATTGGTGGCCCTCTCTGCGCAAGGCCATCATTCGCG-----CTTTGTTATCATC 379
Qy 698 CysCysPheValMetGly-----PheGlyProIleProAsnIleLeuCysAlaGluIle 715
Db 380 TTTGTTGGAGTCTATGAGTGCATATGGGCGGCGGTCATGATTTACTCGGAGAGATC 439
Qy 716 PheProThrArg 719
Db 440 TTCCCCCTTGAAG 451

RESULT 13
US-10-950-009-1142
; Sequence 1142, Application US/10950009
; GENERAL INFORMATION:
; APPLICANT: BERKA, Randy
; APPLICANT: BASHKIROVA, Elena
; APPLICANT: REY, Michael
; TITLE OF INVENTION: Methods For Monitoring Gene Expression
; FILE REFERENCE: 10541.200-US
; CURRENT APPLICATION NUMBER: US/10/950,009
; PRIOR FILING DATE: 2004-09-24
; PRIOR APPLICATION NUMBER: 60/506,140
; NUMBER OF SEQ ID NOS: 1190
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1142
; LENGTH: 491
; TYPE: DNA
; ORGANISM: Trichoderma reesei
US-10-950-009-1142
Alignment Scores:
Pred. No.: 0.00209 Length: 491
Score: 119.00 Matches: 31
Percent Similarity: 50.50% Conservative: 20
Best Local Similarity: 30.69% Mismatches: 46
Query Match: 2.92% Indels: 4
DB: 6 Gaps: 2

US-10-051-909-32 (1-800) x US-10-950-009-1142 (1-491)
Qy 700 PheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArg 719
Db 24 TTTGGCTACAGTTGGGACCTTGGCGCTGCTATCTCATCGTCTGAGATCTGGCCCTTGAGC 83
Qy 720 ValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleVal 739
Db 84 TCTCGACCTTATGTGTCTCTCGGAGCTTCCGCAACTGGTGAACAACTTCATCATC 143
Qy 740 ThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyr 759
Db 144 GGACAGTGACTCCGACATGTTGCAAGGCATC---ACCTATGGCACCATTATCTCTGTTTC 200

```

QY 760 AlaValValCysLeuLeuSerPheValPheLeuLeuValProGluThrLysGly 779
 Db 201 GGAATTATAACATACCTTGGTCCCGCTTGTCTACTTCTTCTCCCGGAGACCAAGCGT 260
 QY 780 MetProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLys 799
 Db 261 CTCACCTCGAA-----GAGATGGACATCATCTTTGGCAGCGAGGCGCCCGCT 311
 QY 800 Ala 800
 Db 312 GCC 314
 RESULT 14
 US-10-632-694A-4/c
 ; Sequence 4, Application US/10632694A
 ; GENERAL INFORMATION:
 ; APPLICANT: Allison, Anthony
 ; TITLE OF INVENTION: MODIFIED ANNEXIN PROTEINS AND METHODS FOR TREATING VASO-OCCLUSIVE
 ; FILE OF INVENTION: SICKLE-CELL DISEASE
 ; FILE REFERENCE: SURR.113
 ; CURRENT APPLICATION NUMBER: US/10/632,694A
 ; CURRENT FILING DATE: 2003-08-01
 ; PRIOR APPLICATION NUMBER: 60/400,718
 ; PRIOR FILING DATE: 2002-08-02
 ; PRIOR APPLICATION NUMBER: 10/080,370
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 4
 ; LENGTH: 2016
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: primer
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (45)..(45)
 ; OTHER INFORMATION: n = a, c, t, or g
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1000)..(1002)
 ; OTHER INFORMATION: n = a, c, t, or g
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1051)..(1053)
 ; OTHER INFORMATION: n = a, c, t, or g
 US-10-632-694A-4
 Alignment Scores:
 Pred. No.: 0.023 Length: 2016
 Score: 118.00 Matches: 158
 Percent Similarity: 33.87% Conservative: 75
 Best Local Similarity: 22.97% Mismatches: 239
 Query Match: 218 Indels: 6
 DB: 31 Gaps: 31
 US-10-051-909-32 (1-800) x US-10-632-694A-4 (1-2016)
 QY 36 ProProAlaSerCysSerSerGln-----GluProValThrSerAspIleLeu 52
 Db 1997 CCGCGGACACAGCAGCAGCGCTTCTTGTAGTCGCGGAGGTGTCGCCCTTGCATCATG 1938
 QY 53 GluAspLysMetSerGlyAlaValLeuValAlaIleValAlaSerIleGly-----Asn 70
 Db 1937 GAGTACAGGGA-GGTGGCGAAGTCTTTCGCGAACTCTCTCGGATGTGTAACAGGTCGAT 1879
 QY 71 LeuLeuGlnGlyTrpAspAsnAla-----ThrIleAlaAlaValLeuTrp 86
 Db 1878 CTCGAGCGGACACAGATCAGCGGATCAGGGTGTGTCGTGCGTGGCGCGCCCTTCAT 1819
 QY 86 ----- 86
 Db 1818 GGCAGTACAGGCTCTCGCCAGGTAGCGGGGATGAGCGGATGACTTCACACGGC 1759

QY 87 IleLysLysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeuIleValSerMet 106
 Db 1758 CAGCAGCAGGTTCTCCAGGTTCGCGAGGTCTCCGGTTCGATGCTCTCTCGATCTGAA 1699
 QY 107 SerLeuIleGlyAlaThrIleVal-ThrThrPheSerGlyProLeuSerAspSerIleG1 126
 Db 1698 GCCGAGATGTCATGTCATGTCGAAACACGCGGAGGTGGGACACGAGCGGCTGCC 1639
 QY 126 YArgArgProMetLeuIleLeuSerSerIleLeuTrpPhePheSerGlyLeuIleMetLe 146
 Db 1638 CAGGATGTGTGATGAACCTTCTCTCGTGGTGCCCACTTCAGTCG-----CCGGC 1588
 QY 146 uTrpSerProAsnValTrpValLeuLeuAlaAraPheValAspGlyPheGlyIleG1 166
 Db 1587 CTGGAACAG-GGCCTGGCGGTCCAGCTCCACCTGGCGCTGTCGATGCG--GGTGTGG 1532
 QY 166 YLeuAlaValThrLeu----- 171
 Db 1531 GGTCCGGTTCGCTGCGAGCAGCACACCAGCATGCGCTGTGTAGTACCGGAGGTGTCG 1472
 QY 172 -----ValProLeuTrpIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAs 189
 Db 1471 CCACCAGTCTCTCTCCAGGTTCGAGGCGGTACTC-----CTCCTCGTAGGCTGCTGA 1418
 QY 189 nThrLeuProGlnPheSerGlyGlyMetPheLeuSerTrpCysMetValPheG1 209
 Db 1417 TGGCGCAGCTCTCTCGGGGTGGGAGCGGATCTCGGTGAGCACCTTCTGTCGG 1358
 QY 209 YMetSerLeuSerProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSe 229
 Db 1357 TCCCGGC-----GCCAGCTTGGCGTGTCTTCAGCTCGTAGG 1322
 QY 229 rLeuPhePheGlyLeuThrIlePheTrpLeuProLysProArgTrpLeuValSe 249
 Db 1321 C-----GTCGTACA 1313
 QY 249 rLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLys----- 266
 Db 1312 GCGCGGAGGCTTTCATCAGGCGCAGCATCAGCTTCTCGAATTGCCGTCAGTCGACT 1253
 QY 267 -----AspAspValSerGlyGluLeuSerLeuLeuGluGly----- 279
 Db 1252 TCATGTCGTTCCACAGGTTCGCGCGAACAGGGTCTTGAACCTCTCGCGCATCTGCTGGC 1193
 QY 280 -LeuGluValGly---GlyAspThrSerIleGluGluTrpIle----- 292
 Db 1192 GCTGGCGCTTGGAGCGCGCGTCAGCATGATGATGCTCTGTCGTCGCCAGGC 1133
 QY 293 -----IleGlyProAlaThrGluAlaAlaAspAspLe 303
 Db 1132 CCTTCATGCGCTTGGCAGCAGCTCGCGTCGCGCGCGCTCGAAGCGGAGAA----- 1078
 QY 303 uValThrAspGly-AspLysGluGlnIleThrLeuTrpGlyProGluGluGlnSerT 323
 Db 1077 -GTCGTCACGGTCCGCGCGAGGCGNNAGTACTTCCGGGCGCC-----T 1034
 QY 323 rp-IleAlaAraProSerLysGlyProIleMetLeuGlySer-ValLeuSerLeuAlaSe 342
 Db 1033 GGAACAGCACCTCCA-----GGCCGATCGAGATCTNNGTCTGCTCGCGCGCGCACA 980
 QY 342 rArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPh 362
 Db 979 GCAGCAGCAGGCGCTT-----CTGTAGTCGCGCG 950
 QY 362 edgYerValHisGluAsnMetProGlnAlaGlySerMetArgSerThrLeuPhePr 382
 Db 949 AGGTGTGCGCTTGTATCATGAGTACAGGAGGTGGCGAAGTCTTTCGGAACTCTTCTGC 890
 QY 382 oAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspG1 402
 Db 889 G----- 889

Db 1357 TCGCCGC-----GCCCCAGCTTGGCGTCTTCAGCTCGTAGG 1322
Qy 229 rLeuPhePheGlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSe 249
Db 1321 C-----GTCGTACA 1313
Qy 249 rLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLys- 266
Db 1312 GCGCGAGGGCTTCATCAGGGCCACCATCAGCTTCTCGAACTTGGCGTTCAGCTCGGACT 1253
Qy 267 -----AspAspValSerGlyGluLeuSerLeuLeuGluGly- 279
Db 1252 TCATGTCGTTACACAGTCCGCGCCGACAGGGTCTTGAACCTCTCGGCGATCTGCTGCG 1193
Qy 280 -LeuGluValGly- -GlyAspThrSerIleGluGluTyrIle- 292
Db 1192 GCTGGCGTTGGAGCGGCGCGTTCAGCAGGTTCAGGATGAGAGTCTCGTGGTCCCGCAGCG 1133
Qy 293 -----IleGlyProAlaThrGluAlaAlaAspLeu 303
Db 1132 CCTTCATGGCTTCGCGCAGCACCTCGCGTCCGCGCGCGCTCGAAGCCGGAGAA----- 1078
Qy 303 uValThrAspGly-AspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerT 323
Db 1077 -GTCGTCACGGTCCGCGCGCAGGCGNNAGTACTTCCGGGGCC-----T 1034
Qy 323 ip-IleAlaArgProSerLysGlyProIleMetLeuGlySer-ValLeuSerLeuAlaSe 342
Db 1033 GGAACAGCACCTCCA-----GCCCCGATCGAGATCTNNNGTGTCTCTCGCGCGGCACA 980
Qy 342 rArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPh 362
Db 979 GCAGCAGCAGGGCTT-----CTGTAGTCGCCGG 950
Qy 362 eGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPhePr 382
Db 949 AGGTGTCGCCCTTCATCGAGTACAGGAGGTGGCGAGTCTTCCGGAACCTCCTTGC 890
Qy 382 oAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGlnTrpAspGl 402
Db 889 G----- 889
Qy 402 uGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyValAlaGlyAspTyrGl 422
Db 888 -----GATGTTGAACAGGTTCGATCTCGAGCGGACACGATCACGC 848
Qy 422 uAspAsn-----LeuHisSerProLeuLeuSerArgGlnAl 434
Db 847 GATCAGGGTGTGTCGTCTCGTCCGCGCGCCCTTCATGGCGTAGTACAGGCTCTCGGCCA 788
Qy 434 aThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMe 454
Db 787 GGTAGCGGGGATGGAGCGATGGACTTCACACGGC----- 751
Qy 454 tArgArgGlnThrLeuLeuGlyGluGlyGly-----AspGlyValSerSerThrAs 471
Db 750 -CAGCAGCAGGTTCTCAGGTTCCGAGAGTCTCGCGGTCTCGATGGTCTCTCGATCTGGA 692
Qy 471 pIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGluAsnGlyAr 491
Db 691 AGCCGAGAGTGTCTATGTTCTGTC-----GAACAGCGCGCGAGGTGGGACA 644
Qy 491 gLysGluGlyGlyPhelysArgValTyrLeuHisGlnGluGlyValPro----- 507
Db 643 CGGAGCGGGTCCCGAGGATGGTGAAGATTCCTCTCGTCCGTGCCCCACTTCAGTCCGC 584
Qy 508 -GlySerArgArgGlySerIleValSerLeu-----ProGlyGlyGlyAs 522
Db 583 CGGCTGGAACAGAGGCTCGGCGCTCCAGCTCCACCTGGGGGTCTCTCGATGGCGGTGTCGG 524
Qy 522 pValPheGluGlySerGluPheValHisAlaAlaLeuValSerGln----- 538

Db 523 GGTCCGCGTTGGCTCGCAGCAGCACCACAGCATGCGCTGGTAGTACCCGAGGTGTCCG 464
Qy 539 -----SerAlaLeuPheSerLysGlyLeuAlaGluProAr 550
Db 463 CCACACAGTCTGCTCTCCAGGTTGGAGCGGTACTCTCTCTCGTAGCGCTGCTT----- 412
Qy 550 gMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLy 570
Db 411 -----GATGGCGCGCAGCTC-----CTCGGGGTGCGGAGCGCA 377
Qy 570 sAspLeuPheGluProGlyValArgAlaLeuValGlyValGlyIleGlnIle-- 589
Db 376 TGATCTC-----GGTCAGCACCTTCTCTCGGTGCCGCGGCCAGCTGG 332
Qy 590 -LeuGlnGlnPheAlaGlyIle-----AsnGlyValLeuTyrTyrThrProGlnIl 606
Db 331 CGTGTCTTCAGCTCTGAGCGGTCTGACAGCGGGAGGCTTCATCAGGGGCCACGATCAGCT 272
Qy 606 eLeuGluGlnAlaGly 611
Db 271 TCTCGAACTTCCCGGT 256

Search completed: October 13, 2004, 23:34:49
Job time : 333.921 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 11:33:22 ; Search time 21.286 Seconds
(without alignments)
3616.147 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRGSLVAVQTPTPLDRR.....PLEVITEFFAVGAKQAQAAKA 800

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2584	63.4	729	2 T06127	probable sugar tra
2	2278	55.9	734	2 H86340	sugar transporter
3	2245	55.1	729	2 T45780	sugar transporter
4	2183	53.6	734	2 T51139	sugar transporter
5	587.5	14.4	457	2 E70070	metabolite transpo
6	536	13.2	580	2 D86426	hypothetical prote
7	533	13.1	461	2 D70073	metabolite transpo
8	524.5	12.9	580	2 D84772	probable sugar tra
9	520.5	12.8	582	2 F71431	hypothetical prote
10	518.5	12.7	469	2 H97054	probable sugar-pro
11	505	12.4	473	2 G69789	sugar transporter
12	493.5	12.1	508	2 G84544	probable sugar tra
13	493	12.1	493	2 A85433	sugar transporter
14	489	12.0	521	2 G84884	probable membrane
15	488	12.0	549	2 T14066	probable sugar tra
16	485.5	11.9	482	2 B69803	metabolite transpo
17	474	11.6	547	2 C84593	probable sugar tra
18	473.5	11.6	464	2 P69587	L-arabinose transp
19	465	11.4	419	2 E69888	metabolite transpo
20	462	11.3	612	2 B40538	myo-inositol trans
21	460.5	11.3	511	2 A45377	probable sugar tra
22	459.5	11.3	511	2 H84536	probable sugar tra
23	458	11.2	560	2 T51485	sugar transporter
24	451	11.1	464	2 F65079	galactose-proton s
25	448	11.0	464	2 C91106	galactose-proton s
26	448	11.0	464	2 F85951	galactose-proton s
27	445	10.9	464	2 A00877	galactose-proton s
28	435.5	10.7	471	2 A50868	L-arabinose isomer
29	434.5	10.7	472	2 S47089	arabinose-proton s

RESULT 1

T06127
Probable sugar transport protein F23E12.140 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C;Accession: T06127
R;Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Hoheisel, submitted to the Protein Sequence Database, April 1999
A;Reference number: Z15485
A;Accession: T06127
A;Molecule type: DNA
A;Residues: 1-729 <BEV>
A;Cross-references: UNIPROT:O65497; EMBL:AL022604; GSPDB:GN00062; ATSP:F23E12.140
A;Experimental source: cultivar Columbia; BAC clone F23E12
C;Genetics:
A;Gene: ATSP:F23E12.140
A;Map position: 4
A;Introns: 27/1; 216/1; 433/1; 519/3; 582/2
C;Keywords: sugar transport; transport protein

Query Match 63.4%; Score 2584; DB 2; Length 729;
Best Local Similarity 68.8%; Pred. No. 3e-172;
Matches 512; Conservative 89; Mismatches 123; Indels 20; Gaps 8;

QY	56	MSGAVLVAIVASIGNLQGWNDATIAAAYLYIKKEFQIQNEPTVEGLIVSKSLGATIVT	115
DB	1	MSGAVLVAIAAAGVNLQGWNDATAGAVLYIKKEFNLESNPSEGLIVAMSLIGATLIT	60
QY	116	TFSGPLSDSIGRRPMLILSSILYFTFSGITMLSPNVYVILLARFVDGFGIGLAVTLVPLY	175
DB	61	TCSGGVADWLGRRPMLILSSILYFVGSVLMSPNVYVILLGRLLDGFVGLVTLVPIY	120
QY	176	ISEIAPSEIRGLLNTLPQSGGGMFLSYCMVFGVSLSPSPDWIRIMLVLAIPSPFFGL	235
DB	121	ISETAPPEIRGLLNTLPQFTSGGMFLSYCMVFGMSLMPSPSWRLMLGVLPSPVFFPL	180
QY	236	TFYIPESPRMLVSKGRMAEAKVLQKRGKDDVSGELSLLEGLVGGDTSEIEVIIIGP	295
DB	181	TVFFLPESPRMLVSKGRMLEAKRVLQRLRGREDVSGEVALLEGIGGTEIEVIIIGP	240
QY	296	ATEAADDLVTDGDKQITLYGPEEQSWIARPSKGPIMLGSVLSLASRHGS-MVANSVPL	354
DB	241	ADEVTDHDIADVDKQIKLYGAEEGLSWARPVKG---GSTMSVLSRHGSTMRSRQSL	296
QY	355	MDPIVTLFGSVHNMPQAGGSVRSLTFNFGSMFSVTDQAHAKNEQWDENLHRDDEEYAS	414
DB	297	IDPLVTLFGSVHVKRMPDT-GSMRSALFPHFGSMFVGGNQRPHEWDENLVGEDEDYPS	355
QY	415	DGAGGGEYNLHSPLLSROATGAEGKDI VHHGHRGSALSMMRQTLIGEGGDVSSDTIGG	474
DB	356	D-HGDDSEDDLSPILSRQTTSMN-KDMPTAHGTLISIFRHGSOVQGAQAGAGSMGIGG	413

QY 475 GWLAWKWSKEGNGRKEGKFRVYHQBVGSGRRGSIUSLPGGSDVPEGSEFVHAAA 534
 DB 414 GWQVANKWTEREDSGOKE-----EGFGSGRRGSIUSLPGGSDGTGE-ADFFVQASA 462
 QY 535 LVQSALFSKGLAEPMSDAAMVHPSSVAAKSGRWKDLFPGVRRALLVGVGQILQOQA 594
 DB 463 LVSPALYSKDLLEKHETIGAMVHPSE-TTKGSIWHLDPGVKRALVGVGQILQOQS 521
 QY 595 GINGVLYTTPQILEQAGAVILSKFGLSSASILISLTLLMLPCIGFAMLMDSGR 654
 DB 522 GINGVLYTTPQILEQAGVIGLLSNMGISSSASILISALTTFVMLPAIAVANRLMDSGR 581
 QY 655 RFLLGCTIPILIASLVLVNSLIDLTGTLAHLSTVSUVVYFCFFVMGFGPFPNLCAR 714
 DB 582 RTLLLTTPILIASLVLVNSLIDLTGTLAHLSTVSUVVYFCFFVMGFGPFPNLCAR 641
 QY 715 IFPRVRGLCTAICAFTHWIGDIIIVTSLPVLNAIGLAGVFSIYAVVCLISFVFLKV 774
 DB 642 IFPRVRGLCTAICAFTHWIGDIIIVTSLPVLNAIGLAGVFSIYAVVCLISFVFLKV 701
 QY 775 PETKGMPLVITTEFFAVGAQAAA 738
 DB 702 PETKGMPLVITTEFFAVGAQAAA 725
 RESULT 2
 Sugar transporter [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: H86340
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.P.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: H86340
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-734 <STO>
 A:Cross-references: UNIPROT:Q9SQ3; GB:AE005172; NID:g4836905; PIDN:AAD30608.1; GSPDB:GN
 A:Map position: 1
 Query Match 55.9%; Score 2278; DB 2; Length 734;
 Best Local Similarity 61.2%; Pred. No. 6.7e-151;
 Matches 462; Conservative 112; Mismatches 145; Indels 36; Gaps 14;
 QY 56 MSGAVLVAIVASTGNLQGDNDATIAAAVLYIKKEFQIQNEPT-VEGLIVNSLIGATIV 114
 DB 1 MKGATLVALAATGNFQGDNDATIAAGVYINKNLNL---PTSVQGLVWVMSLIGATVI 57
 QY 115 TFSGPLSDSISGRPMILSSILYFFSGILMLSPNVYVILLARFVDFGIGLAVTLVPL 174
 DB 58 TTCSGPTSDMLGRPMILSSVVMYFVGLIMLSPNVYVLCFARLNGFGAGLAVTLVPL 117
 QY 175 YISEIAPSEIRGLNLTLPQFSGSGGMFLSYCMVFGMSLSPSPDWRIMLVLAIPSLFFFG 234
 DB 118 YISEIAPSEIRGLNLTLPQFSGSGGMFLSYCMVFTVSLSDSPSWRAMLVGLVSLPILYL 177
 QY 235 LTIFYPLESPRLVSKGRMAAKVQLKRGKDDVSGELSLLEGEVGGDTISIEYIIG 294
 DB 178 LTIFYPLESPRLVSKGRVDEAKRVLQQLCGREDVTGKALLVEGLDGGEXTMEDLIYT 237
 QY 295 PATEAADLVDGKKE-QITLYGPEGQSWIARPSKGPIMLGVSLSIARHGMVNSQVSP 353
 DB 238 LEDHEGDDTLTVDGQMLYGTENQSYARFVPEQ---NSSSLGRSRHGLANQSMI 294

QY 354 LMDPIVTLFGSVHENMPQAGSGMERSTLFPNFGSPFVTDQ--HAKNQWD---BENLHRD 408
 DB 295 LKDPVLNLFGLHFKMPEAGGNTRSGIPPHFGSNFSTTADAPHGKFAHWKXDIESHYNKD 354
 QY 409 DEEVASDGAGDYED---NLHSPLLSRQATGAEKGDIVHGHGHSALSMSRQTLTLLGGGD 465
 DB 355 NDDVATDDGAGDDDDSDNDRSLPMSRQTTSMYD-KDMIPIHPTSGSTLSMRHSTLMQ-GN 412
 QY 466 GVSSTDIQGGQWLAKWSEKGEKGRKGGPKRYLHQEGVPGSGRRGSIUSLPGGSDVFE 525
 DB 413 GESSMGIGGGVHMGVRYENDE-----YKRYLKEDGAE-SRRGSIUSIPGPD--G 460
 QY 526 GSEFVHAALYSQSALFSKGLAEPMSDAAMVHPSEVAAGSRWKLFEQVREALLVGV 585
 DB 461 GGSYTHASALYSRSLVLPKPS-----VHGSAMVPEKTAASGPLMSALPECPVKKALVGV 515
 QY 586 GIQLIQQFAGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLLMLPCIGFA 645
 DB 516 GIQLIQQFSGINGVLYTTPQILERAGVDILLSSLSLSSISASPLISGLTLLMLPAIWA 575
 QY 646 MLMDSGRRELLGTTPILIASLVLVNSLIDLTGTLAHLSTVSUVVYFCFFVMGFG 705
 DB 576 MRLMDVSGRRSLLTTPVLIVSLVLSLISLHISKVYNAALSTGCVLYFCFFVMGFG 635
 QY 706 PIPNLCIAIEPTVRGLCTAICAFTHWIGDIIIVTSLPVLNAIGLAGVFSIYAVVCL 765
 DB 636 PIPNLCIAIEPTVRGLCTAICAFTHWIGDIIIVTSLPVLNAIGLAGVFSIYAVVCL 695
 QY 766 SFVFFVLKVPETKGMPLVITTEFFAVGAQAAA 800
 DB 696 SWIFVYMKVPETKGMPLVITTEFFAVGAQAAA 729
 RESULT 3
 T45780
 sugar transporter-like protein - Arabidopsis thaliana
 N:Alternate names: protein F26013.130
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
 C:Accession: T45780
 R:Peiseny, M.; Berger, C.; Cooke, R.; Grellet, F.; Laudie, M.; Mewes, H.W.; Lemcke, K.;
 submitted to the Protein Sequence Database, December 1999
 A:Reference number: 223013
 A:Accession: T45780
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-729
 A:Cross-references: UNIPROT:Q9SD00; EMBL:AL133452
 A:Experimental source: cultivar Columbia; BAC clone F26013
 C:Genetics:
 A:Map position: 3
 A:Introns: 27/1; 216/1; 523/3; 578/2
 A:Note: F26013.130
 Query Match 55.1%; Score 2245; DB 2; Length 729;
 Best Local Similarity 60.4%; Pred. No. 1.3e-148;
 Matches 457; Conservative 110; Mismatches 144; Indels 46; Gaps 14;
 QY 56 MSGAVLVAIVASTGNLQGDNDATIAAAVLYIKKEFQIQNEPT-VEGLIVNSLIGATIV 115
 DB 1 MRSVTVLALAAAGNMLQGDNDATIAAGVYIKKEFQIQNEPT-VEGLIVNSLIGATIT 60
 QY 116 TFSGPLSDSISGRPMILSSILYFFSGILMLSPNVYVILLARFVDFGIGLAVTLVPL 175
 DB 61 TFSGPLSDSISGRPMILSSILYFFSGILMLSPNVYVILLARFVDFGIGLAVTLVPL 120
 QY 176 ISETAPSEIRGLNLTLPQFSGSGGMFLSYCMVFGMSLSPSPDWRIMLVLAIPSLFFFG 235
 DB 121 ISETAPSEIRGLNLTLPQFSGSGGMFLSYCMVFGMSLSPSPDWRIMLVLAIPSLFFFG 180
 QY 236 TIFYPLESPRLVSKGRMAAKVQLKRGKDDVSGELSLLEGEVGGDTISIEYIIG 295

Db 181 AAFELPESPRWLVSKGRMDKARVQLQRLRGREDVSGELALLVGLGVGKDTSTIEEVIGP 240
QY 296 ATEAADLVLDGDKKE-QITLYGPEEGOSWIARSKGPIMLGSLVSLASRHSWMVQSVPLM 355
Db 241 DNEENEGCNELPKDQIKLYGPEEGOSWMAKPVKQG-----SSLALASRHSGLMPLGSGSLM 296
QY 356 DPTVTLFGSVHEHMP--QAGGSRSTLFPNFGSMFSTVDQHAKEQWDEENLHRDDEEYA 413
Db 297 DPLVTLFGSVHEHMP--QAGGSRSTLFPNFGSMFSTVDQHAKEQWDEENLHRDDEEYA 413
QY 414 SDGAGGSDYEDNLHSPLLSRQATGAEGKDIIVHGHRSALSMMRQTL-LGEGGSDVSSTDI 472
Db 350 -----SSQDENLNSPLSPQT--EPDD--YHQTGTVMHRRQSSLPFMANVGETATATSI 401
QY 473 GGGHQLAWKSEKEGNGRK-EGGFRKVLKQF-----GVPGSRGSIIVSL-POGGDV 523
Db 402 GGGHQLAWKYNKVDGCKRWGLQRMYYTHEETANNNTNIPFSRRGSLLSHPHEDGHH 461
QY 524 FEGSEFVHAALYSQSALFSGKLAEPMSDAAMVHPSEVAAGSRWKDLPFEPGVRRALLV 583
Db 462 DVNGYVYOAALYSQASMPGKGGETAML-----PKEV-KDQPGWRLEKEFGVKREALMV 514
QY 584 GVGIILOQFAGINGVLYYTPQILEQAGVAVILSKFGLSSASASIISSITLTLMLPCIG 643
Db 515 GVGIILOQFAGINGVLYYTPQILEETGVSSLLTNLGSASASLLISALTTLMLPCI- 573
QY 644 FAMILMDSGRRLFLIGTIPILIASILVIVSNLIDGLTAHALLSTVSVIVVFCFVMVG 703
Db 574 -----LVMSRLMSTIIPILISLTLVIGSLNGLGSGINALISTASVTVYLSFVMG 626
QY 704 FGPINILCAIEIPFVRGICIAICAFTFWIGDIIVTVYSLPVMNLAIAGLGVFSIYAVVC 763
Db 627 FGAIPNILCSIEIPFVRGICIAICAFTFWIGDIIVTVYSLPVMNLAIAGLGVFSIYAVVC 763
QY 764 LISFVFLKVPETKGMPLVITEFFAVGAKQAQAA 800
Db 687 AVAWVFLKVPETKGMPLVITEFFAVGAKQAQAA 723

RESULT 4
T51139
sugar transport protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 28-Jul-2000 #sequence_revision 28-Jul-2000 #text_change 09-Jul-2004
C:Accession: T51139
R:Tjaden, J.; Neuhaus, E.
A:Submitted to the EMBL Data Library, August 1995
A:Description: A new sugar transport protein from Arabidopsis thaliana.
A:Reference number: 225311
A:Accession: T51139
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-734 <TUA>
A:Cross-references: UNIPROT:Q96290; EMBL:Z50752; PIDN:CAA90628.1

Query Match 53.6%; Score 2183; DB 2; Length 734;
Best Local Similarity 59.5%; Pred. No. 2.9e-144;
Matches 449; Conservative 110; Mismatches 160; Indels 36; Gaps 14;

QY 56 MSGAVLVAIVASIGMLQGDWNTAIAAALVYIKKEFOLQNEPT-VEGLIVSMSLIGATIV 114
Db 1 MKGATVLAALATIGNFQGDWNTAIAAMVYINKDNLN---PTSVQGLVVAWSLIGATVI 57
QY 115 TTFSGPLSDSITGRPMILISILYFFSGGLIMLSPNYYVLLARFVDGFGIGLAVTLVPL 174
Db 58 TTCGSPISDMLGRPMILISILYFFSGGLIMLSPNYYVLLARFVDGFGIGLAVTLVPL 117
QY 175 YISEIAPSEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPSPDWIRMLGVLAIPSLFFFG 234
Db 118 YISTAPPEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPSPDWIRMLGVLAIPSLFFFG 234
QY 235 LTIFLPEPRWLVSKGRMAEAKVQLKRGKQDVSGELSLLEGLGVGKDTSTIEEYIG 294

Db 178 LTVFVLPESPRWLVSKGRMDKARVQLQRLRGREDVTDQHAKEQWDEENLHRDDEEYA 237
QY 295 PATEAADLVLDGDKKE-QITLYGPEEGOSWIARSKGPIMLGSLVSLASRHSWMVQSVPLM 353
Db 238 LEDHEGDDTLETVDGQIRLYGTHENOSYLARVPVEQ---NSSLGLSRHSGLANQSMI 294
QY 354 LMDPVTTLFGSVHEHMP--QAGGSRSTLFPNFGSMFSTVDQHAKEQWDEENLHRDDEEYA 408
Db 295 LMDPVTTLFGSVHEHMP--QAGGSRSTLFPNFGSMFSTVDQHAKEQWDEENLHRDDEEYA 408
QY 409 DEEYASDAGGSDYED---NLHSPLLSRQATGAEGKDIIVHGHRSALSMMRQTL-LGEGGSD 465
Db 355 NDDYATDDGAGDDDDSDNLASPLMSRQTTSMQ-KDMIHPHTSGTSLSMRRHSTLMQ-GN 412
QY 466 GVSSTIDGGHQLAWKSEKEGNGRK-EGGFRKVLKQF-----GVPGSRGSIIVSL-POGGDV 523
Db 413 GSSWGIIGGHHMGVRYENDE-----YRYLKEDGAE-SRRGSIIVSGPDP--G 460
QY 526 GSEFVHAALYSQSALFSGKLAEPMSDAAMVHPSEVAAGSRWKDLPFEPGVRRALLV 583
Db 461 GGSYIHASALYSRSLVGFPS-----VHGSAMVPEKIAASGFLWSALLEPGVKRALVGV 515
QY 586 GVGIILOQFAGINGVLYYTPQILEQAGVAVILSKFGLSSASASIISSITLTLMLPCIG 643
Db 516 GVGIILOQFAGINGVLYYTPQILERAGVDLISLSSGLSSISASFLISGTTTLMLPAIYVA 575
QY 646 MLLMDSGRRLFLIGTIPILIASILVIVSNLIDGLTAHALLSTVSVIVVFCFVMVG 703
Db 576 MELMDVSGERSLLWTIPVILVSVIVSELHISKVNNALSTGCVLVFCFVMVG 635
QY 706 PIPNILCAIEIPFVRGICIAICAFTFWIGDIIVTVYSLPVMNLAIAGLGVFSIYAVVC 765
Db 636 PPQTSVLKSSQQAQDREGICIAICAFTFWIGDIIVTVYSLPVMNLAIAGLGVFSIYAVVC 695
QY 766 SFVFLKVPETKGMPLVITEFFAVGAKQAQAA 800
Db 696 SWIFVFLKVPETKGMPLVITEFFAVGAKQAQAA 729

RESULT 5
E70070
metabolite transport protein homolog ywtG - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: E70070
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrati, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallucci, M.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadie, J.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, Y.; Sekowska, A.; Seror, A.; Authors: Yoshikawa, H.F.; Yamamoto, H.; Yamane, K.; Yasunori, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunori, A.; Tosato, V.; Uchiyama, A.; Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: E70070
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-457 <KUN>
A:Cross-references: UNIPROT:P96742; GB:Z99122; GB:AL009126; NID:G2636029; PIDN:CAB15600.1
A:Experimental source: strain 168
C:Genetics:
A:Gene: ywtG
C:Superfamily: glucose transport protein

Query Match 14.4%; Score 587.5; DB 2; Length 457;
Best Local Similarity 22.8%; Pred. No. 3e-33;

Matches 168; Conservative 102; Mismatches 162; Indels 305; Gaps 12;	
Qy	55 KMSGAVLVAIVASIGNLQGDWNTAAAVLYIKKEFQLONEPTVEGLIVMSLIGATIV 114
Db	2 KKQSNWLYPFGALGALYGVDTGVISGAILFMKKELGL--NAPTEGLVSSLLVGAAIL 59
Qy	115 TTFSGPLSDSGIRPMLLSILYFFSGGLIMLSPNVYLLARFVDGFGIGLAVTLVPL 174
Db	60 SGAAGKUTDFGRKKAIVAAALFCIGGLGVALAPNTGVMLFRILGLAVGTSITIVPL 119
Qy	175 YISIAPESEIRGLINTLPQSGGGMFLSYCMVFGMSLSPDWRMLGLVLAIPSPFFFG 234
Db	120 YLSLAPKPKRGALSSNLQMLTITVILLSY--IVNYIFADAERWMLGLAAPSLL-- 176
Qy	235 LTIPYLPESPRWLVSGRMAEAKVLOKRGKDDVSGELSLLEGLLEVGDTSIEYIIG 294
Db	177 IGLFMPESPRLFTNGEESKAKILEKLRGTDI----- 211
Qy	295 PATEAADLVTDGKQITLYGPEEGSWIARPSKGPIMLGSLASRHGSMVNSQVPL 354
Db	212 ----- 211
Qy	355 MDPIVTLFGSVHENMPQAGGSMRSTLPNFGSMFVTDQAKNEQWDEENLHRDDEYAS 414
Db	212 -----DOE--- 214
Qy	415 DGAGDYEDNLHSPLLSRQATGAEGKDIVHGHGSGALSRRQTLLEGGDGVSSDTIG 474
Db	215 -----IH-----DI-- 218
Qy	475 GWQLAKWSEKEGNGRKEGFKRVYLHQEGVPGSRGSIIVSLPGGDVPEGSEFVHAAA 534
Db	219 -----KEAEKQDEGGLKE----- 231
Qy	535 LVQSALFSGKLAPRSDAAMVHPSEVAAKGRWKDLPEPGVRRALLVGVGIQILQOFA 594
Db	232 -----LFDPWVRPALIAGLGLAFLOQFT 254
Qy	595 GINGVLYTTPQILEQAGVAVILSKFGLSSASAIL-----ISSLTLLMLPCIGFAMLLMD 650
Db	255 GTNTIIYAKFTNVG-----FG--NSALIGTVGIGTVNVLMTL-----VAIKID 300
Qy	651 LSGRRFLLGTPIPIIASVILVSNLIDLGTLAHLLSTVSVIVFPCFVMGFGPIPMI 710
Db	301 KIGRKPLLLFCNAGWVSLIVLALVNLFFNTTAASTVVICLGFTIVFVAVSWGVVWV 360
Qy	711 LCABIFPVRVGLCIACTFWDIGDIIVTSLVPMNAIGLAGVFSIYAVVCLISFVVF 770
Db	361 MLPFLPLVVRGIGTGVSTLMLHVGTLIVSLTYPILMEALIGSYLIIYAIGIMAFLV 420
Qy	771 FLKVPETKGMPLVITE 787
Db	421 RFKVETKGRSLBIEQ 437

RESULT 6

De6426
 hypothetical protein F12P21.2 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: D86426
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: D86426

A>Status: preliminary	
A:Molecule type: DNA	
A:Residues: 1-580 <STO>	
A:Cross-references: UNIPROT:Q9C757; GB:AB005172; NID:g11120774; PIDN:AAG30955.1; GSPDB:G	
C:Genetics:	
A:Map position: 1	
C:Superfamily: glucose transport protein	
Query Match 13.2%; Score 536; DB 2; Length 580;	
Best Local Similarity 21.6%; Pred. No. 1.6e-29;	
Matches 175; Conservative 93; Mismatches 169; Indels 372; Gaps 11;	
Qy	62 VAIVASIGNLQGDWNTAAAVLYIKKEFQLONEPTVEGLIVMSLIGATIVTTSFGP 120
Db	31 LAFSAGIGGLFGYDVTGVISGAILYIRDDFKSDRNTWLQEMIVSMVAVAGAIAGG 90
Qy	121 LSDSGIRPMLLSILYFFSGGLIMLSPNVYLLARFVDGFGIGLAVTLVPLIYSEIA 180
Db	91 ANDKLGRSAILMADFLLLGAIIMAAAPNPSLLVGVGVFVGLGVGMASMTAPLYISEAS 150
Qy	181 PSEIRGLINTLPQSGGGMFLSYCMVFGMSLSPDWRMLGLVLAIPSPFFFGTLTIFYL 240
Db	151 PAKIRGALVSTNGFLITGGQFLSYLINLAFT-DVGTWRWMLGIAGIPALLOFLV-MFTL 208
Qy	241 PESPRWLVSGRMAEAKVLOKRGKDDVSGELSLLEGLLEVGDTSIEYIIGPATEAA 300
Db	209 PESPRWLVRKREBEAKAILRRIYSABEDVEQEIIRALKDSVET----- 250
Qy	301 DDLVTDGKQITLYGPEEGSWIARPSKGPIMLGSLVSLASRHGSMVNSQVPLMDPIVT 360
Db	251 -EILEGSSSEKINMI----- 264
Qy	361 LFGSVHENMPQAGGSMRSTLPNFGSMFVTDQAKNEQWDEENLHRDDEYASDGAGD 420
Db	265 ----- 264
Qy	421 YEDNLHSPLLSRQATGAEGKDIVHGHGSGALSRRQTLLEGGDGVSSDTIGGGWQLAW 480
Db	265 -----KLCRAKT----- 271
Qy	481 KWSEKEGNGRKEGFKRVYLHQEGVPGSRGSIIVSLPGGDVPEGSEFVHAAALVQSQA 540
Db	272 ----- 271
Qy	541 LFSKGLAEPMSDAMVHPSEVAAKGRWKDLPEPGVRRALLVGVGIQILQOFAINGVL 600
Db	272 -----VERGLIAGVGLQVFOQFYGINVTM 295
Qy	601 YVTQILEQAGVAVILSKFGLSSASAILSSLTLLMLPCIGFAMLLMDLSGRFPLLG 660
Db	296 YISPTIVQLAGFA-----SNRTALLSLVTAGLNAPGSIISIYIFIDRIGKLLII 346
Qy	661 TIPILIASLVIL-----VVSNLIDIGT----- 682
Db	347 SLFGVIISLIGLTGVFYEAATHAPAISSLETQRENNISCPDYKSAMNTNADWCMTCLKAS 406
Qy	683 -----LAH-----ALLSTSVIVY 696
Db	407 SPSCGYCSSPGIKHEHPGACWISDSDVKDLCHENRMLWYTRGCPNSFGWFAALLGLGLIIF 466
Qy	697 FCCFVMGFGPIPNILCAIFPVRVGLCIACTFWDIGDIIVTSLVPMNAIGLAGV 756
Db	467 FS---PGMGTVPIVNSEIYPLRFRGICGGIAATANWISNLIVAQSFLSLTEAIGTSWTF 523
Qy	757 SIYAVVCLISFVVFVFLKVPETKGMPLV 785
Db	524 LIFGVISVIALLFVWVCVPETKGMPLV 552

RESULT 7

D70073
 metabolite transport protein homolog yxoc - Bacillus subtilis
 C:Species: Bacillus subtilis

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: D70073
R:Kunze, P.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertoni, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallenbach, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y. M.; Ogawa, K.; Ogilwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. A:Authors: Yoshikawa, H.F.; Zumsstein, E.; Yoshikawa, H.; Danchin, A. Bacillus subtilis. A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis. A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: D70073
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-461 <KUN>
A:Cross-references: UNIPROT:P46333; GB:299124; GB:AL009126; NID:G2636442; PIDN:CAB16017. A:Experimental source: strain 168
C:Genetics:
C:Gene: YxcC
C:Superfamily: glucose transport protein
Query Match 13.1%; Score 533; DB 2; Length 461;
Best Local Similarity 21.6%; Pred. No. 1.9e-29;
Matches 157; Conservative 102; Mismatches 168; Indels 300; Gaps 12;
QY 61 LVAVASIGNLLOQWNNATIAAIVYKKEFQNEPTVEGLIVSMILGATIVTFSGP 120
DB 9 MIYFFGALGGLLYGDTGVISGALLFINNDIPLTT-LTBGLVSMLLGAIFGSALSGT 66
QY 121 LPSIGRRPMLIISSILYFFSGGLIMWPNVYLLARFVDFGIGLAVTLVPLYSIA 180
DB 67 CSDRWGRKRVFVLSIIFFIICALAFSQTIGMLASRVILGAVGGSTALVEVYLSMA 126
QY 181 PSIRGLLNTLPQSGGGMFLSYCMVFGMSLSPDPDRIMLVLAIPSPFFGLTIFYL 240
DB 127 PTKIRGLTGMNNLMIVTGILLAY--LVNLTFFPFAWRWVGLAIPAVLL-IGIAFM 183
QY 241 PESPRMLVSKGRMAEAKVQLKRGKDDVSGELSLLEGLEVGDDTSIEEYIIGPATEAA 300
DB 184 PESPRMLVSKGRSBEARRINITHDPKDIEMELAEKMQ----- 222
QY 301 DDLVTDGKEQITLYGPBGQSWIARPSKGPIMLSGLASRHSVMNQSVPLMDPIVT 360
DB 223 -----EAEKRETL----- 231
QY 361 LFGSVHNMPPQAGGSMRSTLFPNFGSMFVTDQAKNEQWDEENLHRDDEYASDGAGD 420
DB 232 ----- 231
QY 421 YEDNLHSLSRQATGAEGKDIVHHGRGALSMMRQTLLEGGDGVSTDDIGGGQLAW 480
DB 232 -----GV-----LKA 236
QY 481 KWEKEGNGRKEGGFKRVYLHQEGVPGSRGSIIVSLPGGCVFEGSEFVHAALVSQA 540
DB 237 KW----- 238
QY 541 LFSKGLAEPRMSDAAMVHPSEVAAGSRWKDLPEFGVRALLVGVGTIQLQFAGINGVL 600
DB 239 -----IRPMLLIGVGLAIFQQAAGINTVI 262
QY 601 YYPQILQCAQAVILSKFGSSASAILLSLTLLMLPCIGFAMLLMDLSGRRELL- 659
DB 263 YYAP-----TIFKAGLGTSASALGTWIGILNIMCI-TAMLILDRVGRKKLLIW 312
QY 660 GTIPILLIASIVLVSNLIDLGTLAH-ALLSTVSVIVYFCFVMGFGFIPNILLCAEIFT 718

DB 313 GSVGITLSLAISGV--LLTGLSASTAMTVVFLGVVIVFYQATGPGVWVLMPELPPS 370
QY 719 RVRGCTCAICAFITWIGDIIIVTYSPLWMLNAGLAGVFSIYAVVCLISFVFLKVPETK 778
DB 371 KARGAATGFTTLVLVSAANLIVSLVFFPLMSANGIAWFMVFSVCLISLFFFAFYVNPETK 430
QY 779 GMPLEVI 785
DB 431 GKSLEEI 437
RESULT 8
DB4772
probable sugar transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
A:Accession: D84772
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana. A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84772
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-580 <STO>
A:Cross-references: UNIPROT:Q9ZQP6; GB:AE002093; NID:G4263781; PIDN:AAD15441.1; GSPDB:GN C:Genetics:
C:Gene: At2g35740
C:Superfamily: glucose transport protein
Query Match 12.9%; Score 524.5; DB 2; Length 580;
Best Local Similarity 20.9%; Pred. No. 1e-28;
Matches 171; Conservative 102; Mismatches 175; Indels 369; Gaps 13;
QY 62 VAIVASIGNLLOQWNNATIAAIVYKKEF-QLQNEPTVEGLIVSMILGATIVTFSGP 120
DB 29 LALSAGIGLLFGYNTGVIAGALLYKEEFGEVDNKTLQEIIVSMTVAGAIVGAIGW 88
QY 121 LSDSISGRPMLIISILYFFSGGLIMWPNVYLLARFVDFGIGLAVTLVPLYSIA 180
DB 89 YNDKFGRRMSVLIADVLFLGALVNVIAHAPVILGLIIVFGVGMASMTSPLYISEMS 148
QY 181 PSIRGLLNTLPQSGGGMFLSYCMVFGMSLSPDRIMLVLAIPSPFFGLTIFYL 240
DB 149 PARIRGALVSTNGLLITGGQFLSYLINLAFVHTPG-TWRWMLGVSAPAIQFCL-MLTL 206
QY 241 PESPRMLVSKGRMAEAKVQLKRGKDDVSGELSLLEGLEVGDDTSIEEYIIGPATEAA 300
DB 207 PESPRMLVSKGRSBEARRINITHDPKDIEMELAEKSVR-----AETA 251
QY 301 DDLVTDGKEQITLYGPBGQSWIARPSKGPIMLSGLASRHSVMNQSVPLMDPIVT 360
DB 252 DE----- 253
QY 361 LFGSVHNMPPQAGGSMRSTLFPNFGSMFVTDQAKNEQWDEENLHRDDEYASDGAGD 420
DB 254 ----- 253
QY 421 YEDNLHSLSRQATGAEGKDIVHHGRGALSMMRQTLLEGGDGVSTDDIGGGQLAW 480
DB 254 -----DLIGH----- 258
QY 481 KWEKEGNGRKEGGFKRVYLHQEGVPGSRGSIIVSLPGGCVFEGSEFVHAALVSQA 540
DB 259 TFSDK-----LRGA----- 267
QY 541 LFSKGLAEPRMSDAAMVHPSEVAAGSRWKDLPEFGVRALLVGVGTIQLQFAGINGVL 600
DB 268 -----LSNPFVRHGLAAGITVQAQFVGINTVM 296

QY 601 YPTQILEQAGVAVILSKFGLSSASASILSSITLMLPCIG--FAMLLMDLGRFLL 658
 DB 297 YSPTILOFAGYA-----SNKTAMALALITS-----GLNAVGVVSMFVDYGRKLM 345
 QY 659 LGTIPILIASILV-----673
 DB 346 IISMFGIITCLVLAAVFNEASNAHPKIDKDSRNPFAKNATCPAPAPFTASRPPSNWNC 405
 QY 674 -----VSNLIDLTIAHA-----LLSTVSVIV 695
 DB 406 MKCLQYDCGFCNSGAQYAPACIVOSADMKALCHSKGRFTFFXDGCPKFGYLAIVFLGL 465
 QY 696 YFCFVNGFGPIPNILCABIPFTRVGLCAICAPTFWIGDIIVTYSLPVMLNAIGLAGV 755
 DB 466 YIIVAFMGTVPWVNSEIYPLRYRLAGGIAAVNWSNLVVSFTFLTNVAGSGT 525
 QY 756 FSIYAVVCLISFVFLKVPETKGMPLVITEFFAVG 792
 DB 526 FLLPAGSSAVGLFFIWLVPETKGLQFEVEKLEGG 562
 RESULT 9
 F71431
 C:Species: Arabidopsis thaliana
 A:Variety: Columbia
 C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 09-Jul-2004
 C:Accession: F71431
 R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirx
 P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terryn, N.; Giel
 avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
 Nature 391, 485-488, 1998
 A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
 erhoft, A.; Moeres, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans
 C.; Chlutzis, N.
 A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
 A:Reference number: A71400; MUID:98121113; PMID:9461215
 A:Accession: F71431
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-582 <BEV>
 A:Cross-references: UNIPROT:O23492; GB:297341; NID:G2244991; PID:G2245004
 C:Genetics:
 A:Map position: 4COP9-4C3845
 C:Superfamily: glucose transport protein
 Query Match 12.8%; Score 520.5; DB 2; Length 582;
 Best Local Similarity 20.3%; Pred. No. 1.9e-28;
 Matches 166; Conservative 106; Mismatches 180; Indels 365; Gaps 12;
 QY 62 VAIVASIGNLQGDWNTAAAVLYIKKF-QLONEPTVEGLIVSMISLIGATVITTFSGP 120
 DB 30 LALSAGITGGLFGYDTCVISGALLFIKEDFEDVKKTKLQSTIVSMAVAGAVGAAGW 89
 QY 121 LPSIGRRPMLILSSILYFSGILMLWSNVVYLLARFVDGFGICLAVTLVPLVISEA 180
 DB 90 INDFGRMSILLADVLFIGALVMAFAPAPVIVGRFVGGMASMTSPVISEAS 149
 QY 181 PSBIRGLNTLPQFSGSGMFLSCYCMVFGMSLSPSPDRIMLGLVLAIPSLFFGLTIFYL 240
 DB 150 PARIRGALVSTNGLLITGQFTSYLINLAFVHTFG-TWRWMLGVAGVPAIVQFVL-MLSL 207
 QY 241 PESPRVLVSGRMAEAKVKVLOKRGDDYSGELSLLEGLVGGDTSTEEYIIGATEAA 300
 DB 208 PESPRVLYRKORIAESRAILERIYPADVEAEAKLUSVE-----AEKA 252
 QY 301 DLIVTDGDEKQITLYGPESGQSIARPSKGPIMGLSVLSASRHGSMVNSQSVPLMDPIVT 360
 DB 253 DEAI-----256
 QY 361 LFGSVHENPQAGGSRSTLFPNFGSFKSVTDQHAKEQWDEENLHRDDEYASDAGGD 420

DB 257 -----IGDSFSA-----263
 QY 421 YEDNLHSPLSRQATGAGKDIVHGHRSALSMMROTLLGEGDGVSTDDIGGWOLAW 480
 DB 264 -----KLKGAFG-----270
 QY 481 KWSKEGEGNGRKGKRVYLLHQSGVFGSRGSIYSLPGGSDVFECSFVHAAALVQSA 540
 DB 271 -----270
 QY 541 LFSKGLAEPRMSDAAMVHPSEVAAKGRWKDLFPQVRRALLVGVGIQLQQFAGINGVL 600
 DB 271 -----NPVVRGLAAGITVQVAAQFVGINTVM 297
 QY 601 YPTQILEQAGVAVILSKFGLSSASASILSSITLMLPCIGFAMLLMDLGRFLLG 660
 DB 298 YSPTILOFAGYA-----SNKTAMALALITS-----VSMFVDYGRKLMII 348
 QY 661 TIPILIASLVIL-----VVSNLIDL-----680
 DB 349 SMFGIILACLIILATVFSQAAIHAPKIDAFESRTFAPNATCSAYAPLAENAPPSRWNCMK 408
 QY 681 -----GTIAHA-----LLSTVSVIVYF 697
 DB 409 CLRSECGFCASGVQPYAPGACVWLSDMKATCSRGRTFFKDGCPKFGFLAIVFLGLYI 468
 QY 698 CCFVNGFGPIPNILCABIPFTRVGLCAICAPTFWIGDIIVTYSLPVMLNAIGLAGVFS 757
 DB 469 VYVAFMGTVPWVNSEIYPLRYRLAGGIAAVNWSNLVVSFTFLTNVAGSGTFL 528
 QY 758 IYAVVCLISFVFLKVPETKGMPLVITEFFAVGAK 794
 DB 529 LFAGSTIGLFFIWLVPETKGLQFEVEKLEGG 565
 RESULT 10
 H97064
 probable sugar-proton symporter [imported] - Clostridium acetobutylicum
 C:Species: Clostridium acetobutylicum
 C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
 C:Accession: H97064
 R:Nolling, J.; Breston, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
 J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4838, 2001
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
 A:Reference number: A96900; MUID:21359325; PMID:21359325
 A:Accession: H97064
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-469 <KUR>
 A:Cross-references: UNIPROT:Q97JE7; GB:AE001437; PID:AAK79307.1; PID:GL5024271; GSPDB:GN
 A:Experimental source: Clostridium acetobutylicum ATCC824
 C:Genetics:
 A:Gene: CAC1339
 C:Superfamily: glucose transport protein
 Query Match 12.7%; Score 518.5; DB 2; Length 469;
 Best Local Similarity 21.8%; Pred. No. 2e-28;
 Matches 164; Conservative 94; Mismatches 175; Indels 319; Gaps 12;
 QY 60 VLVAIVASIGNLQGDWNTAAAVLYIKKFQLONEPTVEGLIVSMISLIGATVITTFSG 119
 DB 15 VLISCAAGLGLLYGYDTAVISCAIGFLKLYNL--SPAMQGFVSIIMVGGVLGVGFSG 72
 QY 120 PLSDSIGRRPMLILSSILYFSGILMLWSNVVYLLARFVDGFGICLAVTLVPLVISEI 179
 DB 73 FLGDAIGRRVLMALAFSAIVSISISTSAFMLIFARIVGGLGIGMASALSVYITEC 132
 QY 180 APSEIRGLNTLPQFSGSGMFLSCYCMVFGM-----SLSPSPDRIMLGLVLAIPSLFF 233
 DB 133 APSIRGLSSILQLFTILGISITFFNLGINVMGSETWRVSTGWYMLACGIVPAIVFL 192
 QY 234 GLTIFYLPESPRVLVSGRMAEAKVKVLOKRGDDYSGELSLLEGLVGGDTSTEEYI 293

Db	193	-ITLFPVPSRPLVKSGNIKKAALVLTAKINGAEIAKQLDLSISKSLATENDSSL-----	246
Qy	294	GPATEAADLVDGDKKEQITLYCPGEGSQIARPSKGPIMLSVLSLASHSGSMVNSQVP	353
Db	247	-----GQ-----	248
Qy	354	LMDPIVTLFGSVHENMPQAGSMRSTLFPNFGSMFVTDQHAKNEQWDEENLHRDDEYA	413
Db	249	-----	248
Qy	414	SDGAGGYEDNLHSPILLSRQATGAEGKDIVHHGRGSALSMMRQTLLEGGGDGVSTDIG	473
Db	249	-----	248
Qy	474	GGWQAWKWEKEGEGKRGKRVYLHQEGVPGSRGSIIVSLPGGGDVFPESGFVHAA	533
Db	249	-----	248
Qy	534	ALVQSQALFSKGLAEPMSDAMVHPSEVAAGSRWKDLFEPQVRRALLVGVGIQLOQF	593
Db	249	-----	248
Qy	594	AGNGVLYTPOILEQAGVAVILSKFGLSSASASIL--ISSLTLLMLPCIGFAMLLMDL	651
Db	271	IGNNSTYIGPEIFQWIGP-----KNSFLATSVIGVEVFSTIL-----AMFLDK	318
Qy	652	SGRRFLL-LGTIPILIASILVILVSNLIDGLTAHALLSTVVI-VYFCFWMGF-----G	705
Db	319	LGRKLMELGSAAMAVFLLI-----GTSFYIKLSNGFVILFIICFVVSFCISMG	369
Qy	706	PIPNILCABEFPFRVGLCAICAFWIGDIIIVTSLPVLNAIGLAGVFESIYAVVCLT	765
Db	370	PIPWIMLPEIFPNHLRARATGATIFLWGANWAGQFTPMLLNGIGGATYFIWFCGINVI	429
Qy	766	SFVFLKVPETKGMPLVITEFFAVGAKQAA 797	
Db	430	CFLVITKVPETKNSLEIEKEWIPKQNA 461	
RESULT 11			
G69789			
sugar transporter homolog ydjk - Bacillus subtilis			
C:Species: Bacillus subtilis			
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004			
C:Accession: G69789			
A:Authors: F. Ogasawara, N.; Meszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berton, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997			
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.P. Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mausel, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Fortetelle Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schreuter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seron akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. T.; Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A. Bacillus subtilis. A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis. A:Reference number: A69580; MUID:98044033; PMID:9384377			
A:Accession: G69789			
A:Status: preliminary; nucleic acid sequence not shown; translation not shown			
A:Molecule type: DNA			
A:Residues: 1-473 <UN>			
A:Cross-references: UNIPROT:O34718; GB:Z99107; GB:AL009126; NID:G2632866; PIDN:CAB12442.			
A:Experimental source: strain 168			
C:Genetics:			
A:Gene: ydjk			
C:Superfamily: glucose transport protein			
Query Match 12.48; Score 505; DB 2; Length 473;			

Best Local Similarity 22.3%; Pred. No. 1.8e-27; Matches 168; Conservative 106; Mismatches 176; Indels 302; Gaps 16;			
Qy	54	DKMGAVLVAIVASIGNLQGDWNTAAAVLYIKKEFQIQNEPTVEGLIVSMELIGATI	113
Db	6	NQMSFLATILVSTFGGLFGYDTGVNLGALPYWGEDQLNLNAFTTEGLVTSLLFGAAL	65
Qy	114	VTTSGPLSDSISGRPPMLILSSILYFFSGLIMLSPNYYVLLARFVDGFGIGLAVTLVP	173
Db	66	GAVFGGMSDFNGRRKNILFLAVIFFSTIGCTFAPNTVMISFRVLGAVGASVTVP	125
Qy	174	LYIEIAPSEIRGLNLTLPQSGGGMFLSCM--VEGMSLSP--SPDWRIMLGLVLAIPSL	230
Db	126	AYLAEVSVESGRMTQNELMIVSGQLLAFVFNAILGTTMGDNHVRFMVLVATSLPAL	185
Qy	231	F-FGLTIFYLPESPRMLVSKGRMAEAKVKQLURGKDDVSGELSLLEGLVEGDDTSIE	289
Db	186	FLFTCM--IRMPESPRMLVSKGRKEDALRVLKIR	218
Qy	290	EYIIGPATEAADLVDGDKKEQITLYGPEQSQIARPSKGPIMLSVLSLASHSGSMWN	349
Db	219	-----	218
Qy	350	QSVPLMDPIVTLFGSVHENMPQAGSMRSTLFPNFGSMFVTDQHAKNEQWDEENLHRD	409
Db	219	-----	219
Qy	410	BEYASDGAGGYEDNLHSPILLSRQATGAEGKDIVHHGRGSALSMMRQTLLEGGGDVSS	469
Db	220	EKRAA-----AEQEI	230
Qy	470	TDIGGWQAWKWEKEGEGKRGKRVYLHQEGVPGSRGSIIVSLPGGGDVFESEF	529
Db	231	-----BFAFK	236
Qy	530	VHAAALVSQALFSKGLAEPMSDAMVHPSEVAAGSRWKDLFEPQVRRALLVGVGIQI	589
Db	237	-----EDQLEKAT-----FKDLSVPWVRVIFVIGLGA	265
Qy	590	LQPFAGINGVLYTPOILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLLM	649
Db	286	VQQTGVSIMYGVTEILRNSG---FQTEAALIGNIANGVISVLATFV-----GIWLL	315
Qy	650	DLSGRR-FLDGTIPILIASILVILVSNLIDGLTAHALLSTVSVIVVFCFWMG-FQPI	707
Db	316	GRVGRPPMLMTGLTGTALLIGIFSLVLE-GSPALPYV-VLSLTVTFLAFOQGAISPV	373
Qy	708	PNILCABEFPFRVGLCAICAFWIGDIIIVTSLPVLNAIGLAGVFESIYAVVCLISF	767
Db	374	TWMLSEIFPLRLRGLGNGVTVFCLMMVNFVAFSTFPILLAAIGLSTTFVIFVGLGICSV	433
Qy	768	VFVFLKVPETKGMPLVITEFFAV-----GAKQ 795	
Db	434	LFVKEFLPETKGLSLEQLEENFRAYDHSQAKK 465	
RESULT 12			
G84564			
probable sugar transporter [imported] - Arabidopsis thaliana			
C:Species: Arabidopsis thaliana (mouse-ear cress)			
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004			
C:Accession: G84564			
A:Authors: Ribin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; A. M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999			
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.			
A:Reference number: A84420; MUID:20083467; PMID:10617197			
A:Accession: G84564			
A:Status: preliminary			
A:Molecule type: DNA			
A:Residues: 1-508 <STO>			
A:Cross-references: UNIPROT:Q9ZNS0; GB:AE002093; NID:G4218010; PIDN:AAID2218.1; GSPDB:GNC			

C;Genetics:

A;Gene: At2g18480

A;Map position: 2

C;Superfamily: glucose transport protein

Query Match 12.1%; Score 493.5; DB 2; Length 508;

Best Local Similarity 22.7%; Pred. No. 1.2e-26;

Matches 170; Conservative 99; Mismatches 174; Indels 305; Gaps 19;

QY 63 AIVASIGNLQGDWNTAIAAALVYIKKEFQIQNEPTVE---GLIVSMISLIGATIVTTFSG 119
 DB 26 AIVASIIIFGDTGWSGAQIFIRDDKI-NDTQIEVLAGINLCALVGSUT----AG 80
 QY 120 PLSDISGRRLMLILSSILYFSGGLIMLSPNVYLLARFVDGFGIGLAVTLVPLVYISEI 179
 DB 81 KTSDEVIGRRYTIASAVIFLVGSLMGYGNYPVLAVGRCIAGVGVGFALMTAPVYSABI 140
 QY 180 APSEIRGLNTLPQFSGSGMPL---SYCMVFGMSLSPSPDWIRMLGVLAIPSLFF-PG 234
 DB 141 SSASHRGFLTSLPELCTISIGILGYVSNYC--FG-KITLKLGRMLGLIAAFPSGLIAPG 197
 QY 235 LTFIFYPESPRLVSKGRMAEAKVKLQKRGKDDVSGELSLLEGLEVGSDTSIEEYIIG 294
 DB 198 IT--RMPESPRLVMQGRLEAKKIM----- 221
 QY 295 PATEAADLVTDDGKEQITLYGPBEGQSWIARPSKGPIMLGSLVLSLRHSGMWNQSVPL 354
 DB 222 ----- 221
 QY 355 MDPIVTLFGSVHNMPOAGGSMSTLFPNFGSMFSTVDQAKNEQWDEENLRHDEEYAS 414
 DB 222 -----VLVSNTEEBEER----- 234
 QY 415 DGAGGYEDNLHPSLRQATGAEGKDI VHGHGRGALSNNRQTLLEGEGDGVSTDIGG 474
 DB 235 -----FRDIL-----TAAE-----VDVTEKEVGG 254
 QY 475 GWQLAKWMSKEGEGKRGKRVYLHOGVPGSRGSIPLPGGDDVFEGSEFVHAAA 534
 DB 255 G-----VKKNVHK----- 263
 QY 535 LVSQSALFSKGLAEPMSDAAMVHPSEVAAKSGRWKDLF---EPGVRALLVGVGIQILQ 591
 DB 264 -----SVNRELVIKRPVAVRLILIAAVGIHFFE 291
 QY 592 QPAGINGVLYTQILEQAGVA---VILSKGLSSASILISSITLLMLPCIGFAML 647
 DB 292 HATGIBAVLYSPRIKKGAVSKDKLLATVGVGLTKAFFII-----IATF 338
 QY 648 LMDLSGRRLFLTLTIPILIASILVVSNLI--DLGTLAHL--LSTVSVIVVFCFCFVMGF 704
 DB 339 LLDKVGRRKLLTSTGMYFALTSLAVSLTMVQRFGLAWALSLSIVSYAFVAFSIGL 398
 QY 705 GPINILCAEIPTRVRGLCIACTFWIGDIIVTYSPLVPMNAIGLAVFSIYAVVCL 764
 DB 399 GPITWYSSEIFPLRLRQAQASIGAVNRMATVSMFSLTKAITTGGVFFVFAIGAV 458
 QY 765 ISFVFFVLKVPETKGMPLVITEFFAVG 792
 DB 459 AAWFFFFMLPETKGLPLEMEKLFGGG 486

RESULT 13

A85433

C;Species: Arabidopsis thaliana (mouse-ear cross)

C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

C;Accession: A85433

R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring

Nature 402, 769-777, 1999

A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A;Reference number: A85001; MUID:20083488; PMID:10617198

A;Accession: A85433

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-493 <STO>

A;Cross-references: UNIPROT:O23213; GB:NC_001268; NID:g7270615; PIDN:CAB80333.1; GSPDB:G

C;Genetics:

A;Gene: ATg36670

A;Map position: 4

C;Superfamily: glucose transport protein

Query Match 12.1%; Score 493; DB 2; Length 493;

Best Local Similarity 23.2%; Pred. No. 1.3e-26;

Matches 177; Conservative 102; Mismatches 178; Indels 306; Gaps 23;

QY 48 SDDILEKMSG-----AVLVAIVASIGNLQGDWNTAIAAALVYIKKEFQIQNEPTVE--- 100
 DB 2 ADQISGEKPGAWNRFAQLCAIVASIVSIIFGYDTGVMSGAMVFIEDLK-TNDVQIEVLT 60
 QY 101 GLIVSMISLIGATIVTTFSGPLSDISGRRLMLILSSILYFSGGLIMLSPNVYLLARFV 160
 DB 61 GILNLCALVGSLL-----AGRTSDIIGRRYTIIVLASILFMLGSLIMGPNYPVLLSGRCT 116
 QY 161 DFGIGLAVTLVPLVYISEIAPSEIRGLNTLPQFSGSGMFLSYCMVFGMSLSPSP-DWR 219
 DB 117 AGLGVGFALMVAVPYSAEIAATASHRGLLASLPHLCISIGILGYVIVNYPFSLPMHIGWR 176
 QY 220 IMLGVLAIPSLFF-FGUTIFYPESPRLVSKGRMAEAKVKLQKRGKDDVSGELSLLE 278
 DB 177 LMLGIAAVPSLVLAFG--ILKMPESPRLVMQGRLEAKKIL----- 216
 QY 279 GLEVGGTSTIEEYIIGATPATEAADLVTDDGKEQITLYGPBEGQSWIARPSKGPIMLGSLV 338
 DB 217 -----ELVSN-----SPEAE----- 227
 QY 339 SLASRHSGMWNQSVPLMDPIVTLFGSVHNMPOAGGSMSTLFPNFGSMFSTVDQAKNE 398
 DB 228 -----LR----- 229
 QY 399 QWDEENLRHDEEYASDGAGGYEDNLHPSLRQATGAEGK--DIVHHGRGALSMSR 455
 DB 230 -----FQD-----IKAAAGIDPKCVDDV----- 248
 QY 456 RQTLLEGEGDGVSTDIGGWLAKWMSKEGEGKRGKRVYLHOGVPGSRGSI 515
 DB 249 -----KMEGKK-----THGEGV----- 260
 QY 516 SLPGGDDVFEGSEFVHAAALVVSQSALFSKGLAEPMSDAAMVHPSEVAAKSGRWKDLF-- 573
 DB 261 -----WKELIIL 267
 QY 574 -EPGVRALLVGVGIQILQOQFAGINGVLYTQILEQAGVAVILSKFGLSSASILISS 632
 DB 268 PTPAVRRLVLTALGIHFQHASGIEAVLLYGPRIKKAGITT-KDKLFLVTIGVIM-- 323
 QY 633 LTLLMLPCIGFAMLLMDLSGRRLFLTLTIPILIASILV-----VVSNLI DLGTLAHL 687
 DB 324 KTTFFIT-----ATLLDKVGRKLLTTSVGMVIALTMLGFLGTWAQNA--CGKLAWAL 376
 QY 688 -LSTVSVIVVFCFCFVMGFPIPNILCAEIPTRVRGLCIACTFWIGDIIVTYSPLVM 746
 DB 377 VLSIVAYSFVAFSIGLGPITWVYSSEVFPKLRAQASGLGVAVNRMATVSMFSL 436
 QY 747 LNAIGLAVFSIYAVVCLISFVFLKVPETKGMPLVITEFF 789
 DB 437 TSAITGGAFFMFAVAAVAMNFFFLPETKGSLEIEALF 479

RESULT 14

G84864

probable membrane transporter [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cross)

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C;Accession: G84864

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; P

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Niemman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A/Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: G84864

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-521 <STO>

A/Cross-references: UNIPROT:O22848; GB:AE002093; NID:g2289003; PIDN:AA864332.1; GSPDB:GN

C/Genetics:

A/Gene: At2g43330

A/Map position: 2

C/Supfamily: glucose transport protein

Query Match 12.0%; Score 489; DB 2; Length 521;
Best Local Similarity 21.2%; Pred. No. 2.6e-26;
Matches 159; Conservative 98; Mismatches 179; Indels 314; Gaps 13;

QY 62 VAIVASIGNLQGMNDNATIAAALYIKKEPQLQNEPTVEGL-----IVMSL 108
DB 34 LTVTAGIGLLFGYDTGVISCALLYKDDPEVVKQSSFLQVNVSSFTSKLETIVSMAL 93

QY 109 IGATIVTFSGPLSDSGRPMILSILYFFSGLIMLSPNVYVLLARFVDGFGIGLA 168
DB 94 VGAMIGAAGGWINDYGRKKAFLFADVFAAGAIVMAAADPDYVLIISGRLLVGLGVGA 153

QY 169 VTLVPLVYISITAPSEIRGLNLTLPQSGSGMFLSYCMVFGMSLSPDWRIMGLAIP 228
DB 154 SVTAPVYIAEASPEVRGGLVSTNVLMTGGQFLSYLVNSAFTQVFG-TWRWMLGVSGVP 212

QY 229 SLFFPGITIFLPSRWLVSKGRMAFAKVKLOKRGKDDVSGELSLLEGLVGEGDGI 288
DB 213 AVIQFILMLF-MPESPRLFMKRAEAIQVLR----- 245

QY 289 EEVIIGPATEADDLVTGDKQITLYPREGQSWIARPSKPIMLGSLVLSASRHGSMV 348
DB 246 -----EDEI-----DHLS 260

QY 349 NQSVPLMDPIVTLFGSVHNNPOAGGSMRSTLPNFGSMFSVTDQAKNEQWDEENLHRD 408
DB 246 -----TYDISRL----- 252

QY 409 DEEVASDGAGDYEDNLHSLPSRQATGAEKDIIVHHGRGSALSMRQTLLEGEGDGV 468
DB 253 -----EDEI-----DHLS 260

QY 469 STDIGGQWLAWKSEKEGNGRKEGFKRVYLHQEGVPGSRGSIIVSLPGGDVFESE 528
DB 261 AAE-----BEKQKRTVG-----YL-----DVFRSKE 283

QY 529 FVHAALVSQALFSKGLAEPRMSDAAMVHPSEVAAGSRKWKDLFPGYVRALLVGVIGQ 588
DB 284 -----LRLAFLAGLQ 295

QY 589 ILQOFAGINGVLYYTPQILEQAG-----VAVILSKFGLSSASILITSLTLLMLPCIG 643
DB 296 AFQOFTGINVYSPITVQWAGFHSNQLALFLSLIVAAKNAAGTVV----- 342

QY 644 FAMILMDLGRFRLLLTGIPILIASVLVNSLIDIGTLA-----HALLTSTVIVYFCC 699
DB 343 -GIYFIDHCGRKKLALSGFVIISLILSVSFQKQSETSDGGLYGLWALVGLALYIVF 401

QY 700 FVMGFGPIPLCAEIPPTVRGGLCIACAFTFWIGDIIIVTSLPVMLNAIGLAGVFSY 759
DB 402 FAPGMGFPVTNSEIYPOQVRGICGMSATVWISNLIVAOFTLITAEAGTMTFLIL 461

QY 760 AVVCLISFVFLKVPETKGMPLVITEFF 789
DB 462 AGIATVAVIVFVVPETQGLTFSEVEIQ 491

RESULT 15

T14606

probable sugar transport protein 205 - beet

C/Species: Beta vulgaris (beet)

C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C/Accession: T14606; T14617

R/Chou, T.J.; Bush, D.R.

Plant Physiol. 110, 511-520, 1996

A>Title: Molecular cloning, immunochemical localization to the vacuole, and expression in

A/Reference number: Z18131; MUID:96351183; PMID:8742332

A/Accession: T14606

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-549 <CHI>

A/Cross-references: UNIPROT:P93075; EMBL:U64302; NID:g1778092; PID:g1778093

A/Accession: T14617

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-541 'SVQV' <CH2>

A/Cross-references: EMBL:U64903; NID:g1778094; PID:g1778095

C/Genetics:

A/Note: BvCDNA-205; BvCDNA-397

C/Supfamily: glucose transport protein

C/Keywords: sugar transport

Query Match 12.0%; Score 488; DB 2; Length 549;
Best Local Similarity 21.3%; Pred. No. 3.3e-26;
Matches 164; Conservative 113; Mismatches 194; Indels 298; Gaps 19;

QY 41 SSQEPVTSDDILE-----KMSGAVLVAIVASIGNLQGMNDNATIAAALYIKKEP 91
DB 10 SDPPPTASKVIADPDLKPKPKNKAFAFACATLMTSVLLGYDVGMSGAILVKEDW 69

QY 92 QLONEPTVEGLVNSLIGATIVTFSGPLSDSGRPMILSILYFFSGLIMLSPNV 151
DB 70 H1SD--TQGVLVGILNIYCLFGSPAAGRTSDWIGRRRTIVLAGAIFVFGALLMGFATNY 127

QY 152 YVLLIARFVDGFGIGLAVTLVPLVISEIAPSEIRGLNLTLPQSGSGMFLSYCMVFGMS 211
DB 128 AFLWGVRFVTGIVGVYALMIAPVYTAEVSPPASSRGLTSFVEVPINAGILLYISNLAFS 187

QY 212 LSPSP-DWRIMGLVLAIPSLFPFGLTIFLPSRWLVSKGRMAFAKVKLOKRGKDDVS 270
DB 188 SLPTHLSEFRFLGIGAIPEFI-FLAIGVLAPESPRWLVMOGLGDAKVLAR----- 238

QY 271 GELSLLLEGVGGDTSEEVIIIGPATEADDLVTGDKQITLYPREGQSWIARPSKG 330
DB 239 -----TSD-----SPEAQ----- 247

QY 331 PIMLGSLVLSASRHGSMYVQLMDPIVTLFGSVHNNPOAGGSMRSTLPNFGSMFSV 390
DB 248 ----- 247

QY 391 TDQAKNEQWDEENLHRDDEEVASDGAGDYEDNLHSLPSRQATGAEKDIIVHHGRGS 450
DB 248 -----L 248

QY 451 ALSMRQTLLEGEGDGVSTDIGGQWLAWKSEKEGNGRKEGFKRVYLHQEGVPGSR 510
DB 249 RLSEIKQT-----AGIPA-- 261

QY 511 RGSIVSLPGGDVFGSEFVHAALVSQALFSKGLAEPRMSDAAMVHPSEVAAGSRWK 570
DB 262 -----ECDE-----DIYKVEKTKISGNVWK 283

QY 571 DLF--EPGVRRALLVGVIGIQLQOFAGINGVLYYTPQILEQAGVAVILSKFGLSSAS 627
DB 284 ELFFNPTAVRAVRAVIGIHFQOASGIDAVVLSYSPRIFQSAGITNARKQL-----LAT 338

QY 628 ILISLTLTLLMPCITGFAMLMDSGRFLLLTGIPILIASVLVNSLIDIG-----T 682
DB 339 VAVGVVKTILFIL---VATFQDKYGRRPFLITSVGGMIITLITLAMSITVTDHSHHKT 394

QY 683 LAHALLTSTVIVYFCCFVMGFGPIPLCAEIPPTVRV--GLCIACAFTFWIGDIIVT 740

Db 395 WATALCITM-VCAVASFSIGLGPITWVVSSEVFPLRLRAQGTSMGV-AVNRVVSQVISI 452
QY 741 YSLPVMNLNAGVFSIYAVVCLISFVFVFLKVPETKGMPLVITEFF 789
Db 453 FFLP-LSHKITTCGAPFCGIAIIAWFFPLTLPETRGRTLENWHELP 500

Search completed: October 13, 2004, 11:39:09
Job time : 26.286 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 13, 2004, 11:40:19 ; Search time 6494.01 Seconds
(without alignments)
4489.025 Million cell updates/sec

Title: US-10-051-909-32

Perfect score: 4075

Sequence: 1 IRSGSLAVQTPPTDRLRR.....PLEVITEFFAVGAKQAQAAKA 800

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-Q=/cgn2_1/USPTO.spool_h/US10051909/runat_13102004_123336_19917/app_query.fasta_1.1678
-DB=EST -QMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOTALIGN=200 -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US10051909 @CGN 1 1 5263 @runat_13102004_123336_19917 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hrc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gssi:*
9: gb_gssi2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	3027	74.3	2180	3	AY105508 Zea mays
C	2	1399	34.3	911	8 CC423708
3	1382.5	33.9	845	8	BZ723844 PUCFI60TD
C	4	1279	31.4	795	8 BZ989776
5	1268.5	31.1	1135	3	CNS0A3H4
6	1168	28.7	833	6	CA05405
7	1154	28.3	733	6	CA146057
8	1142	28.0	696	6	CA075335
9	1120	27.5	1082	7	CK208781

10	1101	27.0	829	5	BQ752737	BO752737 WHE4118 E
11	1100	27.0	871	6	CA170109	CA170109 SCQSB108
12	1086	26.7	1107	5	BX841312	BX841312 EX841312
13	1066.5	26.2	848	7	CF445664	CF445664 EST682009
14	1063.5	26.1	769	6	CA243478	CA243478 SCQFLA07
15	1055	25.9	994	6	CA067439	CA067439 SCQAD106
16	1034	25.4	636	6	CF018228	CF018228 QBM9a03 X
17	1033.5	25.4	950	7	CK263963	CK263963 EST710041
18	1030	25.3	692	6	CA112602	CA112602 SCQFLB106
19	1022	25.1	631	6	CA252785	CA252785 SCBFL114
20	1018	25.0	787	6	CF182124	CF182124 JG-2 Chin
21	1014.5	24.9	643	6	CA263039	CA263039 SCPLB202
22	1014.5	24.9	858	6	CA207025	CA207025 SCEQSB100
23	1009	24.8	644	6	CA279697	CA279697 SCCFL800
24	1003	24.6	616	6	CA081000	CA081000 SCCCAM200
25	997	24.5	606	6	CA230633	CA230633 SCQFL300
26	990.5	24.3	745	6	CB635485	CB635485 OS1EB150
C	27	981	594	1	AI861088	AI861088 603011H02
28	980	24.0	620	6	CA167255	CA167255 SCQSB100
29	966	23.7	669	6	CA073177	CA073177 SCEPAM105
30	957	23.5	817	8	BZ819234	BZ819234 PUGDY07TB
31	945	23.2	903	6	CF211894	CF211894 CGF100066
32	941.5	23.1	655	6	CA268224	CA268224 SCQRT305
33	939	23.0	701	6	CD879943	CD879943 AZ04.1061
34	937	23.0	827	6	CA207188	CA207188 SCEQSB100
35	935	22.9	801	7	CK201790	CK201790 FGAS01031
36	930	22.8	859	4	BM817360	BM817360 HC106B05
37	928	22.8	749	5	BU040051	BU040051 PP Leao000
38	924	22.7	688	6	CA502649	CA502649 WHE4338 A
39	921.5	22.6	700	5	BQ870435	BQ870435 QGD3A19-Y
40	921.5	22.6	703	6	CA502490	CA502490 WHE4048 B
41	921	22.6	687	6	CD983252	CD983252 G118.123E
42	916	22.5	847	6	CB981615	CB981615 CAB70005
43	914	22.4	736	2	BF108319	BF108319 EI22F22 N
44	909	22.3	572	4	BM325261	BM325261 PIC1 42 E
45	909	22.3	793	6	CA266132	CA266132 SCULET308

ALIGNMENTS

LOCUS	AY105508	2180 bp	mRNA	linear	HTC 16-OCT-2002
DEFINITION	Zea mays PC011453 mRNA sequence.				
ACCESSION	AY105508				
VERSION	AY105508.1				
KEYWORDS	HTC.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
REFERENCE	1 (bases 1 to 2180)				
AUTHORS	Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.				
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of				
JOURNAL	Overgo Probes				
REFERENCE	2 (bases 1 to 2180)				
AUTHORS	Coe,E.H.				
TITLE	Direct Submission				
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of				
COMMENT	Missouri, Columbia, MO 65211, USA				
	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.				
FEATURES	Location/Qualifiers				
source	1..2180				
	/organism="Zea mays"				

```

/mol_type="mRNA"
/db_xref="MaizeDB:636667"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/notes="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 1,02e-294 Length: 2180
Score: 3027.00 Matches: 591
Percent Similarity: 99.66% Conservatives: 2
Best Local Similarity: 99.33% Mismatches: 2
Query Match: 74.28% Indels: 0
DB: 3 Gaps: 0

```

US-10-051-909-32 (1-800) x AY105508 (1-2180)

```

Qy 206 MetValPheGlyMetSerLeuSerProSerProAspTTPargileMetLeuGlyValLeu 225
Db 50 ATGGTGTGGGATGTCCTGCGCCATCACCGGATGGAGAAATATGCTTGGTGGCTC 109
Qy 226 AlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuProGluSerProArg 245
Db 110 GCGATACCTTCATGTTCTTCTTGGTTGGACAAATATTTATCTTCTGATCTCCAGA 169
Qy 246 TrpLeuValSerGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGly 265
Db 170 TGGCTGTTAGCAAGGTCGATGGCAGAGCGCAAAAGGTTGCAAAAGTTACGGTG 229
Qy 266 LysAspValSerGlyGluLeuSerLeuLeuGluGlyLeuGluValGlyGlyAsp 285
Db 230 AAGACCATGCTCAGGTGAATATGCTTCCTCGAAGGTTGGAGGTGGAGGAGAC 289
Qy 286 ThrSerIleGluGluTyrIleLeuGlyProAlaThrGluAlaAlaAspLeuValThr 305
Db 290 ACTTCATTTGAGAGTACATCATTTGACCTGCGACCGAGGACCGCATGCTTTACT 349
Qy 306 AspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAla 325
Db 350 GACGGTATAGGAACAAATCACACTTTATGGGCCCTGAAGAGCCGATGATGCT 409
Qy 326 ArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGly 345
Db 410 CGACCTTCTAAGGAGCCCATCATGCTTGGAGGTGCTTCTCTTGCATCTCGTCATGG 469
Qy 346 SerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerVal 365
Db 470 AGCATGTGAACACAGAGTATACCCCTATGGATCCGATTTGTGACATTTTGTGATGTC 529
Qy 366 HisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGly 385
Db 530 CATGAGAAATATGCTCAAGCTGGAGGAGATGATGAGAGGACCATGTTTCCAAACTTGA 589
Qy 386 SerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeu 405
Db 590 AGTATGTTCACTGTCACAGATCAGCATGCCCAAAATGAGCAGTGGGATGAAGAGATCTT 649
Qy 406 HisArgAspAspGluGluTyrAlaSerAspGlyValAclyGlyAspTyrGluAspAsnLeu 425
Db 650 CATAGGATGACGAGGAGTACCATCTGATGTTGGCAGAGGTGACTATGAGGCAATCTC 709
Qy 426 HisSerProLeuLeuSerArgGlnAlaThrGlyValaGluGlyLysAspIleValHisHis 445
Db 710 CATAGCCCATTTGCTGTCAGGCGAGGCAACAGGTGCGGAGGGAAGACATTTGACCAT 769
Qy 446 GlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyAsp 465

```

```

Db 770 GGTCAACGTCGAAGTGTCTTTGAGCATGAGAAGGCAAGAGCCCTCTTAGGGAGGTGGAGAT 829
Qy 466 GlyValSerSerThrAspIleGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLys 485
Db 830 GGTGTGACGACACTGATATCGTGGGGATGGGAGCTTGTCTTGAATGTGTGAGAAG 889
Qy 486 GluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGly 505
Db 890 GAAGGTGAGATGTGAGAAGAGAGTGGTTTCAAAGAGTCTACTTGCACCAAGAGGA 949
Qy 506 ValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGlu 525
Db 950 GTTCTGCTCAAGAAGGGGCTCAATTTTCACTTCCCGTGGTGGCATGTTCTTTGAG 1009
Qy 526 GlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGly 545
Db 1010 GGTAGTGAGTTTGTACATGCTGCTTGTAGTAGTCACTGAGTCAAGAGGT 1069
Qy 546 LeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLys 565
Db 1070 CTTGCTGAACACGACATGTCAGATGCTCCATGCTTCAACCATCTGAGTAGTGCACAA 1129
Qy 566 GlySerArgTTPlysAspLeuPheGluProGlyValArgArgAlaLeuValGlyVal 585
Db 1130 GGTTCACGTTTGAAGATTTGTTGAACCTGGAGTGGAGGCTGCTGTAGTCGGTGT 1189
Qy 586 GlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGln 605
Db 1190 GGAATTCAGATCTTCAACAGTTCGTAATAAAGGTTCTGTACTATACCCACAA 1249
Qy 606 IleLeuGlnGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSer 625
Db 1250 ATTCTTGACGAGTGGTGGCAGTATTTCTTCCAAATTTGGTCTCAGCTCGCATCA 1309
Qy 626 AlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAla 645
Db 1310 GCATCCATCTTGATCAGTCTCTCACTACCTACTAATGCTTCTGCTTGGCTTTGCC 1369
Qy 646 MetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeu 665
Db 1370 ATGTGCTTATGAGTCTTTCGGAAGAGGTTTGTGCTGAGTACCAATTTCCAACTTG 1429
Qy 666 IleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHis 685
Db 1430 ATAGCATCTTAGTATATCTTGTGTCATCTTATGATTTGGTGTACACATAGCCCAT 1489
Qy 686 AlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGly 705
Db 1490 GCTTTGCTCTCCACCATCAGTGTATGCTCTACTTCTGCTGCTTCTGTTATGGGATTTGGT 1549
Qy 706 ProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIle 725
Db 1550 CCCATCCCAACATTTATGTCAGAGATCTTCCAAACAGGTTCTGTCGCTCTGTATT 1609
Qy 726 AlaIleCysAlaPheThrPheTrpIleGlyAspIleValThrThrSerLeuProVal 745
Db 1610 GCCATTTGCTCTTACATCTGGATCGAGATATCATCTGTCACCTACAGCCTTCTCTGTG 1669
Qy 746 MetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIle 765
Db 1670 ATGCTGAATGCTATGGAGTGGCGGGTGTTCAGCATATATGAGTCTGATGTTGATT 1729
Qy 766 SerPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIle 785
Db 1730 TCCTTTGTGTTGCTCTTCAAGTCCCTGAGACAAAGGGATGCCCTTGAAGTTATT 1789
Qy 786 ThrGluPhePheAlaValGlyAlaLysGlnAlaLysAlaLysAla 800
Db 1790 ACCGAATTCCTTGGAGTGGTGGCAAGAGCGGCTGCAAAAGCC 1834

```

RESULT 2

CC423708/c

LOCUS

CC423708

911 bp

DNA

linear

GSS 19-MAY-2003

```

DEFINITION PUH0949TB ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZMMBTa499I02,
genomic survey sequence.
ACCESSION CC423708
VERSION CC423708.1 GI:30903798
KEYWORDS GSS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 911)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other GSSs: PUH0949TD
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
    source
        1..911
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMMBTa499I02"
            /clone_lib="ZM_0.6_1.0_KB"
            /notes="Vector: PCR4-TOPO; Site:1: EcoRI; 0.6-1.0 kb high
            Cor selected genomic DNA library"
ORIGIN
Alignment Scores:
Pred. No.: 1,71e-130 Length: 911
Score: 1399.00 Matches: 267
Percent Similarity: 98.90% Conservative: 2
Best Local Similarity: 98.16% Mismatches: 3
Query Match: 34.33% Indels: 0
DB: 8 Gaps: 0
US-10-051-909-32 (1-800) x CC423708 (1-911)
Qy 270 SerGlyLeuSerLeuLeuLeuGlyValGlyAspThrSerLeuGlu 289
Db 818 ACAGGTGAATTGTCCTCTCTCGAAGGTTGGAGGTTGGAGGACACTTCCATTGAA 759
Qy 290 GlnTyrLeuLeuGlyProAlaThrGluAlaAspLeuValThrAspGlyAspLys 309
Db 758 GAGTACATCATTTGACCTGCCAGGCGACCGCATGATCTTCTTACTGACGGTGATAAG 699
Qy 310 GluGlnLeuThrLeuTyrGlyProGluGluGlyGlnSerTrpLeuAlaArgProSerLys 329
Db 698 GAACAAATCACACTTTATGCGCTCGAAGAGGCGAGTCATGATGCTCGACCTTCCAAG 639
Qy 330 GlyProLeuMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsn 349
Db 638 GGACCCAGCAGCTGTGAAGTGCCTTCTTCATCTCTGTCATGCGGAGCATGGTGAC 579
Qy 350 GlnSerValProLeuMetAspProLeuValThrLeuPheGlySerValHisGluAsnMet 369
Db 578 CAGAGTGATACCCCTTATGATCCGATGTCACACTTTTGGTAGTGTCCATGAGATATG 519
Qy 370 ProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSer 389
Db 518 CCTCAAGCTGAGGAAGTATGAGAGCAGCATGTTTCCAAACTTTGGAAGTATGTTCAAT 459
Qy 390 ValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAsp 409

```

```

Db 458 GTCACAGATCAGCATGCCAAAATGAGCAGTGGATGAGAGAAATCTTCATAGGATGAC 399
Qy 410 GluGlnTyrAlaSerAspGlyAlaGlyGlyAspTyrGlnAspAsnLeuHisSerProLeu 429
Db 398 GAGGAGTACGATCTGATGTCGAAGAGGTGACTATGAGGACAACTTCCATAGCCCATG 339
Qy 430 LeuSerArgGlnAlaThrGlyAlaGluGlyLysAspLeuValHisGlyHisArgGly 449
Db 338 CTGTCCAGGAGGAGCAACAGGTGCGGAAGGAGGACATTTGCACCATGTCACCGTGA 279
Qy 450 SerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGlyGlyGlyGlyValSerSer 469
Db 278 AGTCTTTTGGCATGAGAGGCAAAAGCTCTTAGGGAGGTTGGAGATGTTGAGGAGC 219
Qy 470 ThrAspLeuGlyGlyTyrGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsn 489
Db 218 ACTGATATCGTGGGAGTGGACCTTCTTGGAAATGTCAGAGAAGGAGGTGAGAAT 159
Qy 490 GlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySer 509
Db 158 GGTAGAAAGGAAGGTGGTTTCAAAAGAGTCTACTTGCAACAGAGGAGTTCCTGGGTCA 99
Qy 510 ArgArgGlySerIleValSerLeuProGlyGlyGlyValPheGluGlySerGluPhe 529
Db 98 AGAAGGGGCTCAATTTTACCTTCCCGTGGTGGCATGTTCTTGAGGAGTAGTGAGTTT 39
Qy 530 ValHisAlaAlaLeuValSerGlnSerAlaLeu 541
Db 38 GTACATGCTGCTGCTTTTAGTAAGTCAGTCAGCACTT 3
RESULT 3
BZ723844
LOCUS BZ723844
DEFINITION BZ723844 845 bp DNA linear GSS 24-FEB-2003
ACCESSION BZ723844
VERSION BZ723844.1 GI:28519500
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 845)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
    source
        1..845
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMMBTa133J23"
            /clone_lib="ZM_0.6_1.0_KB"
            /note="Vector: PCR4-TOPO; Site:1: EcoRI; 0.6-1.0 kb high
            Cor selected genomic DNA library"
ORIGIN
Alignment Scores:
Pred. No.: 7,09e-129 Length: 845
Score: 1382.50 Matches: 274
Percent Similarity: 97.52% Conservative: 1

```

Best Local Similarity: 97.16% Mismatches: 6
 Query Match: 33.93% Indels: 3
 DB: 8 Gaps: 0

US-10-051-909-32 (1-800) x BZ723844 (1-845)

Qy	303	LeuValThrAspGlyAspLysGluGlnThrLeuTyrGlyProGluGluGlnSer	322
Db	3	CTTGTTACTGCGGTATAGGACAAATCACACTTATGGGCTGAGAGCCAGTCA	62
Qy	323	TrpLeuAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSer	342
Db	63	TGGATTCTCGACCTTCAAGGGACCCAGCATGCTTGAAGTGTCTTCTCTTGCACT	122
Qy	343	ArgHisGlySerMetValLeuGlnSerValProLeuMetAspProIleValThrLeuPhe	362
Db	123	CGTCATGGAGCATGGTGAACAGAGTGTACCCCTTATGGATCCGATTGTGACACTTTT	182
Qy	363	GlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPhePro	382
Db	183	GCTAGTCTCCATGAGAATATGCTCAAGCTGGAGGAAGTATGAGGACACATTTGTTTCA	242
Qy	383	AsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGlu	402
Db	243	AACCTTGGAAAGTATGTTCACTGTACAGATCAGCATGCTAAATAATGAGCAGTGGGATGA	302
Qy	403	GluAsnLeuHisArgAspAspGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGlu	422
Db	303	GAGAACTTTCATAGGATGACGAGGAGTACGCATCTGATGTTGTCAGGAGTGTACTATGAG	362
Qy	423	AspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIle	442
Db	363	GACAATCTCCATAGCCCATTTGCTGTCCAGCAGGCAACAGGTGCGGAGGAGGACAT	422
Qy	443	ValHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGlu	462
Db	423	GTGCACCATGTCACCGTGGAGTGTGTTGAGCATGAGAGGCAAGCCCTCTTAGGGAG	482
Qy	463	GlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTyrGlnLeuAlaTrpLysTrp	482
Db	483	GGTGGAGATGTTGTGACGACACTGTATATCGGTGGGGATGCGAGCTTCTTGGAAATGG	542
Qy	483	SerGluLysGluGlyGluLeuGlyArgGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly	502
Db	543	TCAGAGAAGAGAGTGTAGAAATGTGTAAGAGGAGGTGTTTCAAGAGTCTACTTGGAC	602
Qy	503	GlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyAsp	522
Db	603	CAAGAGGGAGTCTCTGGCTCAAGAGGGGCTCAATTGTTTCACTTCCCGGTGTCGCGAT	662
Qy	523	ValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPhe	542
Db	663	GTTCCTTGGAGGTAGTGTGATGTTGTACATGCTGTCTTTAGTAAGTCAGTCAGCATTTC	722
Qy	543	SerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluVal	562
Db	723	TC-AAGGGCTTGTCTGAACA-CGCATGTCAGATGCTGCATGTTTCAACCCATCTGAGTA	780
Qy	563	AlaAlaLysGlySerArg-TrpLysAspLeuPheGluProGlyValArgAlaLeuLe	582
Db	781	GCTGCCAAAGGTTACGTTTGGAAAGATTGTTTGAACCTGAGGTGAGCGTCCCTGTT	840
Qy	582	uVal 583	
Db	841	AGTC 844	

RESULT 4

BZ989776 795 bp DNA linear GSS 25-MAR-2003
 LOCUS PUGDY07TD ZM 0.6.1.0 KB Zea mays genomic clone ZMWBta362B13,
 DEFINITION genomic survey sequence.
 ACCESSION BZ989776
 VERSION BZ989776.1 GI:29229913

KEYWORDS

SOURCE Zea mays
 ORGANISM Zea mays

GSS.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

1 (bases 1 to 795)

AUTHORS

Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
 Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
 Bennetzen,J.

TITLE

Maize Genomics Consortium

JOURNAL

Unpublished (2003)

COMMENT

Other GSSs: PUGDY07TB

Contact: Cathy Whitelaw

TIGR Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: Tg

Class: sheared ends

Location/Qualifiers

1..795

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMWBta362B13"

/clone_lib="ZM 0.6.1.0 KB"

/note="vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high"

Cor selected genomic DNA library"

ORIGIN

Alignment Scores:

Pred. No.: 1.94e-118 Length: 795

Score: 1279.00 Matches: 242

Percent Similarity: 99.19% Conservative: 2

Best Local Similarity: 98.37% Mismatches: 2

Query Match: 31.39% Indels: 0

DB: 8 Gaps: 0

US-10-051-909-32 (1-800) x BZ989776 (1-795)

Qy 270 SerGlyGlnLeuSerLeuLeuGluValGlyGlyAspThrSerIleGlu 289

Db 738 ACAGGTGAATGTCCCTTCTTCGAAGGGTTGGAGGAGACACTTCCATTGAA 679

Qy 290 GluTyrIleIleGlyProAlaThrGluAlaAlaAspLeuValThrAspGlyAspLys 309

Db 678 GAGTACATCATTTGACCTGCCACCGAGGAGCCCATGATCTTGTACTGACGGTGAATAG 619

Qy 310 GluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLys 329

Db 618 GAACAAATCACTTTATGGCCTGAAGAAGGCCAGTCATGCTGCTCCAGCTTCCCAAG 559

Qy 330 GlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsn 349

Db 558 GGACCCAGCAGTCTTGAAAGTGTGCTTCTTTCATCTCTGATGGAGCATGGTGAAC 499

Qy 350 GlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMet 369

Db 498 CAGAGTGTACCCCTTATGGATCCGATTGTGACACTTTTGGTAGTGTCCATGAGAATATG 439

Qy 370 ProGluAlaGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSer 389

Db 438 CCTCAAGCTGGAGGAAGTATGAGGACCATTTTCCAAACTTTTGAAGTATGTTCAGT 379

Qy 390 ValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAsp 409

Db 378 GTCAAGATCAGCATGCCAAAATGAGCAGTGGGATGAAGAGAAATCTTCATAGGGATGAC 319

Qy 410 GluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAsnLeuHisSerProLeu 429

318 GAGGAGTACGATCTGATGTGTGAGGAGTGTACTATGAGGACAAATCTCCATAGCCCATTTG 259

430 LeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGly 449

258 CTCTCAGGAGGACAAAGAGTGGAGGAGGAGGACATTTGTCACCATGTCACCGTGA 199

450 SerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyAspGlyValSerSer 469

198 AGTGCTTGAGCATGAGAGGCAAGCCCTCTTAGGGGAGGAGTGGAGATGCTGTGAGCAG 139

470 ThrAspIleGlyGlyGlyGlyGlnLeuAlaThrLysThrSerGlyLysGluGlyGluAsn 489

138 ACTGATATCGTGGGAGTGGGAGCTTGTCTGGAATGTGTGAGAGGAGGAGTGAAT 79

490 GlyArgLysGluGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 509

78 GGTAGAAAGGAGGAGTGGTTCCTCAAGAGTCTACTTGCACCAAGAGGAGTCTCTAGCTCA 19

510 ArgArgGlySerIleVal 515

18 AGAAGGGGCTCAATTGTT 1

CNS0A3H4 1135 bp mRNA linear HTC 06-FEB-2004

Arabidopsis thaliana Full-length cDNA Complete sequence from clone

GSLSIL50ZE01 of Silique of strain col-0 of Arabidopsis thaliana

(thale cress).

EX828912

Version: BX828912.1 GI:42461191

Keywords: HTC; GSLT cDNA

Source: Arabidopsis thaliana (thale cress)

Organism: Arabidopsis thaliana

Reference: Arabidopsis thaliana

Authors: Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

1 (bases 1 to 1135);

Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissbach,J. and Salanoubat,M.

Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation

Unpublished

2 (bases 1 to 1135)

Genoscope.

Direct Submission

Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

The sequences are based on single pass reads.

Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.

Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissbach J., Salanoubat M.

URGV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.

http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length

http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

Location/Qualifiers

1..1135

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/strain="col-0"

/db_xref="taxon:3702"

/clone="GSLT5IL50ZE01"

/tissue_type="Silique"

/plasmid="PCWVSORT_6"

complement(1..1135)

gene

ORIGIN

Alignment Scores:

Pred. No.: 3,96e-117 Length: 1135

Score: 1268,50 Matches: 247

Percent Similarity: 82,82% Conservative: 47

Best local Similarity: 69,58% Mismatches: 58

Query Match: 31,13% Indels: 3

Gaps: 3

DB:

US-10-051-909-32 (1-800) x CNS0A3H4 (1-1135)

445 HisGlyHisArgGlySerAlaLeuSerMetArgArg---GlnThrLeuLeuGlyGluGly 463

QY

1 CACACTGCTCATGGAACCTTTCTTACCTTCAGACATGGAAGTCAAGTGCAGGAGCTCAA 60

DB

464 GlyAspGlyValSerSerThrAspIleGlyGlyGlyGlyGlnLeuAlaThrLysThrSer 483

QY

61 GGGGAGGAGGCGGTAGTATCGGGATGCGAGTGGTGGCAAGTGGCATGGAATGGACG 120

DB

484 GluLysGluGlyGluAsnGlyArgLysGluGlyGlyGlyGlyGlyGlyGlyGlyGly 503

QY

121 GAAAGAGAAGATGAATCGGAGCAGAAAGAGTGGGTTTAAACGGATATATCTTGATCAA 180

DB

504 GluGlyValProGlySerArgArgGlySerIleValSerLeuP:cglyGlyGlyAspVal 523

QY

181 GAAGTTTCCAGGATCTCGACGTGGCTCAATTGTTTCATTCCTGCTGGTGGTATGGAACC 240

DB

524 PheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSer 543

QY

241 GGTGAG---GCAGATTTGTACAGCGTCTGTTGGTTAGCCACCAAGCTCTTTATTTC 297

DB

544 LysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAla 563

QY

298 AAAGACCTTCTCAAAGAACATACAAATGCTGCTGTATGGTACATCCATCCGAA---ACA 354

DB

564 AlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuVal 583

QY

355 ACTAAGGGTCAATTTGGCATGATCTTCATGATCTCGAGTCAAGCGTGCATTAGTCGTA 414

DB

584 GlyValGlyIleGlnIleGlnGlnPheAlaGlyIleAsnGlyValLeuTrpThr 603

QY

415 GGAGTTGGACTTCAAAATATCTTACAGCAGTCTTCAGGACATCAACGGAGTCTTTTACTACA 474

DB

604 ProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSer 623

QY

475 CCGCAAAATCTTGACAGCGGGTCTCGGATCTCTACTATCGAACATCGGGATTAGTCT 534

DB

624 AlaSerAlaSerIleLeuSerSerLeuThrThrLeuLeuMetLeuProCysIleGly 643

QY

535 TCCTCAGCATCCTTACTTATAGTGCATTGACAACTTTGTGATGTTTACCTGCAATAGCT 594

DB

644 PheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIlePro 663

QY

595 GTTGAATGAGGCTCATGATCTTTCTGGTCAAGGACCTTGTCTCTCACACGATACCA 654

DB

664 IleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeu 683

QY

655 ATCTTGATAGCATCTCTATTGTTTGTAGTAACTCTCAAAATCTTGTTCACATGACAGCAT 714

DB

684 AlaHisAlaLeuLeuSerThrValSerValIleValTrpPheCysPheValMetGly 703

QY

715 GTGCACGGCGTCTTATCAACCGGTGTGCTCTACTTCTGCTTCTTCGTGATGGGT 774

DB

704 PheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeu 723

QY

775 TTCGGTCTGCTCCAAACATCTCTGTTTCAGAGATTTTTCACAGCTCGAGTCCGCGGAATC 834

DB

724 CysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTrpSerLeu 743

QY

835 TGCATCGGCATCTGGCACTCACCTTCTGGATCTGTGACATAATCGTCACTTACAGTCTC 894

DB

JOURNAL
COMMENT

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
 Contact: Arruda P
 Centro de Biologia Molecular e Engenharia Genetica
 Universidade Estadual de Campinas
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil
 Tel: 55 19 3788 1137
 Fax: 55 19 3788 1089
 Email: parruda@unicamp.br
 Cloning distribution: Cloning distribution information can be found
 through the Brazilian Clone Collection Center (BCCC) at
 http://www.bccccenter.fcav.unesp.br
 Plate: 076 row: D column: 10
 Seq primer: T7 Promoter Primer.
 Location/Qualifiers
 1..733
 /organism="Saccharum officinarum"
 /mol_type="mRNA"
 /db_xref="taxon:4547"
 /clone="SCVPR2076D10"
 /lab_host="DH10B"
 /clone_lib="RT2"
 /note="Organ: Root tips(0.3cm-long) from adult plants;
 vector: pSport1; Site: 1: SalI; Site: 2: NotI; An
 unidirectional cDNA library generated from [Root
 tips(0.3cm-long) from adult plants]. cDNA was prepared
 from polyA+ mRNA using SuperScript Plasmid System kit
 (Invitrogen). The double-strand cDNAs were fractionated
 in a sepharose Cl-2B 40cm-columns and fragments sizing
 between 0.8 and 1.5 Kb were directionally cloned into the
 vector. Details of each source of RNA and library
 construction can be obtained at
 http://sucet.lad.ic.unicamp.br/public"

FEATURES

source

421 TGGAAATGGTCAGAGAGAGGTGAGGATGTAAGAGGAGGTGCTCAAAAGAGTC 480
 500 TyrLeuHisGlnGluGlyValProGlySerArgGlySerIleValSerLeuProGly 519
 481 TACTTACACCAAGAGGAGTTCCTGGCTCAAGAATGGCTCAATGTTCTCTCTGT 540
 520 GlyGlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSer 539
 541 GTGGCAATGTTCTTGGAGTGGCGAGTTGTATCATGCTGTCTTAAAGCCGTCGA 600
 540 AlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisPro 559
 601 GCACITTTTCTCGAAGATCTTACCGAACCGCATGCTGGTGTGCCATGTTTCAACCA 660
 560 SerGluValAlaAlaLysGlySerArg-TripLeuAspLeuPheGluProGlyValArg 579
 661 TCCAGGTAGTGTCCAAAGTTCAAGTTGGAAGATTTGTTGAACCTGTGTGAGCGG 720
 579 gAlaLeuLeuVal 583
 721 TGCCCTTGTAGTC 733

RESULT 8
 CA075335
 LOCUS
 DEFINITION
 5', mRNA sequence.
 CA075335 696 bp mRNA linear EST 23-SEP-2003
 SCJFAM1069A06.g AM1 Saccharum officinarum cDNA clone SCJFAM1069A06
 CA075335
 CA075335.1 GI:34927608
 EST.
 Saccharum officinarum
 Saccharum officinarum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
 Complex.
 1 (bases 1 to 696)
 Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
 The libraries that made SUCEST
 Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
 Contact: Arruda P
 Centro de Biologia Molecular e Engenharia Genetica
 Universidade Estadual de Campinas
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil
 Tel: 55 19 3788 1137
 Fax: 55 19 3788 1089
 Email: parruda@unicamp.br
 Cloning distribution: cloning distribution information can be found
 through the Brazilian Clone Collection Center (BCCC) at
 http://www.bccccenter.fcav.unesp.br
 Plate: 069 row: A column: 06
 Seq primer: T7 Promoter Primer.
 Location/Qualifiers
 1..696
 /organism="Saccharum officinarum"
 /mol_type="mRNA"
 /db_xref="taxon:4547"
 /clone="SCJFAM1069A06"
 /lab_host="DH10B"
 /clone_lib="AMI"
 /note="Organ: Apical meristem and tissues surrounding of
 mature plants; Vector: pSport1; Site: 1: SalI; Site: 2:
 NotI; An unidirectional cDNA library generated from
 [Apical meristem and tissues surrounding of mature
 plants]. cDNA was prepared from polyA+ mRNA using
 SuperScript Plasmid System Kit (Invitrogen). The
 double-strand cDNAs were fractionated in a sepharose
 Cl-2B 40cm-columns and fragments sizing between 0.8 and
 1.5 Kb were directionally cloned into the vector. Details
 of each source of RNA and library construction can be
 obtained at http://sucet.lad.ic.unicamp.br/public"

FEATURES
 source
 1..733
 /organism="Saccharum officinarum"
 /mol_type="mRNA"
 /db_xref="taxon:4547"
 /clone="SCVPR2076D10"
 /lab_host="DH10B"
 /clone_lib="RT2"
 /note="Organ: Root tips(0.3cm-long) from adult plants;
 vector: pSport1; Site: 1: SalI; Site: 2: NotI; An
 unidirectional cDNA library generated from [Root
 tips(0.3cm-long) from adult plants]. cDNA was prepared
 from polyA+ mRNA using SuperScript Plasmid System kit
 (Invitrogen). The double-strand cDNAs were fractionated
 in a sepharose Cl-2B 40cm-columns and fragments sizing
 between 0.8 and 1.5 Kb were directionally cloned into the
 vector. Details of each source of RNA and library
 construction can be obtained at
 http://sucet.lad.ic.unicamp.br/public"

ORIGIN

Alignment Scores:
 Pred. No.: 7,71e-106 Length: 733
 Score: 1154.00 Matches: 221
 Percent Similarity: 93.88% Conservative: 9
 Best Local Similarity: 90.20% Mismatches: 14
 Query Match: 28.32% Indels: 1
 DB: 6 Gaps: 0

US-10-051-909-32 (1-800) x CA146057 (1-733)

Qy 340 LeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleVal 359
 1 CTCGCATCTCGTCATGCGCAGCATGTCAGACAGAGTGTACCCCTTATGGATCGATTGTG 60
 Qy 360 ThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThr 379
 61 ACACITTTTGGAGTGTCCACGAGAATATGCTCAAGCTGGAGAGATGATGAGGAGCACA 120
 Qy 380 LeuPheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGln 399
 121 TTGTTTCAAACTTTGAAGTATGTTTCACTGTCACAGATGCGCATGCGCAAAATGAGCAG 180
 Qy 400 TrpAspGluGlnLeuHisArgAspGluGlyTyrAlaSerAspGlyAlaGlyGly 419
 181 TGGGACCAAGAGAATCTTACAGGAGCATGAGGAGTATGTCATCTGATGTTGCGAGGAGGT 240
 Qy 420 AspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGly 439
 241 GATTATGAGGACAATCTCCAGGCCCATTTGCTGCAGGAGCAGACAACAGTGTGGAAAGGA 300
 Qy 440 LysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgGlnThrLeu 459
 301 AAGGACATTTGTCACCATGTTGTCACCGTGAAGTCTTTGAGCATGAGAGCAAGGCTC 360
 Qy 460 LeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyTyrGlnLeuAla 479
 361 TTGGGGAGGCTGGAGAGGTTGAGCAGCACTGATATTGTTGAGGATGCGACCTTGCA 420
 Qy 480 TrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgVal 499

ORIGIN

Alignment Scores:
 Pred. No.: 1.16e-104 Length: 596
 Score: 1142.00 Matches: 214
 Percent Similarity: 95.22% Conservative: 5
 Best Local Similarity: 93.04% Mismatches: 11
 Query Match: 28.02% Indels: 0
 DB: 6 Gaps: 0

US-10-051-909-32 (1-800) x CA075335 (1-696)

QY 305 ThrAspGlyAspLysGluGlnThrLeuTyrGlyProGluGluGlnSerTrpIle 324
 Db 1 ACTGATGTTGATAAGCAACAAATCACACTTATGGCTGAAGAAGCCAGTCATGGATT 60

QY 325 AlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerIleAlaSerArgHis 344
 Db 61 GCTCGACCTTCCAAAGGACCAGCATGCTTGGAGTGTCTTCTCGCATCTCGTCAT 120

QY 345 GlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySer 364
 Db 121 GCCAGCATGGTGAACACAGAGTGTACCCCTTATGGATCCGATGTGCACATTTTGGAGT 180

QY 365 ValHisGluAsnMetProGlnAlaGlySerMetArgSerThrLeuPheProAsnPhe 384
 Db 181 GTCCACAGAAATATGCCTCAAGCTGGAGGAAGTATGAGGAGCAATGTTTCCAAACTTT 240

QY 385 GlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsn 404
 Db 241 GGAAGTATGTCAGTGTCAAGATCAGATCAGATGCCAAATAGCAGTGGGACGAGAGAT 300

QY 405 LeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsn 424
 Db 301 CTTTACAGGACCATGAGGAGTATGATCTGATGCTGTCAGAGTGTATATGAGGACAAT 360

QY 425 LeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHis 444
 Db 361 CTCACAGCCCATGCTGTCTCAGGACGACAAAGTATGGAAGAAAGACATTTGTCAC 420

QY 445 HisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGlyGly 464
 Db 421 CATGGTCACCGTGAAGTGTCTTTCAGCATGAGAAGCAAGCCCTTTGGGGAGGCTGA 480

QY 465 AspGlyValSerSerThrAspIleGlyGlyTyrGlnLeuAlaTrpLysTrpSerGlu 484
 Db 481 GAGGCTGTGAGCAGCACTGATATTTGGTGGAGGATGGCAGCTTCATGGAATGGTCAGAG 540

QY 485 LysGluGlyGluAsnGlyArgLysGluGlyGlyPhelysArgValTyrLeuHisGlnGlu 504
 Db 541 AAGGAAGTTCAGGATGTTAAGAGGAAGTGTGTTTCAAAAGAGTCTACTTACACCAAGAG 600

QY 505 GlyValProGlySerArgGlySerIleValSerLeuProGlyGlyAspValPhe 524
 Db 601 GGAATTCCTGCTCAAGAAATGGGCTCANTGNTTCATCTTCTGTTGGTGGCATGTTCT 660

QY 525 GluGlySerGluPheValHisAlaAla 534
 Db 661 GAGGTTGGCGATGTTGTACATGCTGTGCT 690

RESULT 9
 CK208781
 LOCUS FGAS020500 Triticum aestivum FGAS: Library 5 GATE 7 Triticum
 DEFINITION aestivum cDNA, mRNA sequence.
 ACCESSION CK208781
 VERSION CK208781.1 GI:39571171
 KEYWORDS EST.
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooideae; Triticeae; Triticum.
 REFERENCE 1 (bases 1 to 1082)

AUTHORS Allard, F., Crosby, W.L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D., Genswein, B., Graf, R., Gulick, P., Hrycan, L.D., Laroche, A., Links, M.G., McCarthy, E.B., Monroy, A., Muzak, I., Nilsson, D., Penniket, C., Roach, J.L. and Sarhan, F.
 TITLE Functional Genomics of Abiotic Stress In Wheat and Canola Crops
 JOURNAL Unpublished (2003)
 COMMENT Contact: Wm L Crosby
 Bioinformatics
 University of Saskatchewan, Department of Computer Science
 1C101 Engineering Building, 57 Campus Drive, Saskatoon,
 Saskatchewan, S7N 5A9, Canada
 Tel: 306 966 1769
 Fax: 306 966 2033
 Email: fgas@estscs.usask.ca
 This sequence is the direct result of the Base calling software
 Phred (default parameters). It is the raw base calls. To aid in the
 identification of the high quality insert the software Lucy
 (default parameters) has been run on this sequence. Lucy identified
 the region [31,846].
 Plate: LS8015 row: E column: 19.
 Location/Qualifiers
 1. 1082
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /db_xref="taxon:4565"
 /clone_lib="Triticum aestivum FGAS: Library 5 GATE 7"
 /note="Vector: pQMV.SPORT6; Crown and developmental stages
 of spike formation in wheat cultivar Norstar. 4 mRNA
 populations were combined before constructing the library.
 The first mRNA population is from 1cm crown sections after
 30 days of cold acclimation. The second is from 1cm crown
 sections after 11 days of deacclimation (before
 deacclimation plants were fully vernalized for 49 days).
 The third is from different developmental stages of spike
 formation (5 to 50mm) that still have not emerged from the
 leaf (dissection required). The last is from different
 developmental stages of spike and seed formation after
 having emerged from the leaf (visible). First strand
 synthesis in this library was done in the presence of
 methylated dCTP thereby protecting from internal cleavage
 with NotI."

ORIGIN
 Alignment Scores:
 Pred. No.: 3.98e-102 Length: 1082
 Score: 1120.00 Matches: 242
 Percent Similarity: 81.23% Conservative: 35
 Best Local Similarity: 70.97% Mismatches: 49
 Query Match: 27.48% Indels: 20
 DB: 7 Gaps: 4

US-10-051-909-32 (1-800) x CK208781 (1-1082)

QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuGlnGlyTrp 75
 Db 65 ATGTGCGGCGCGTGTGTGTCGTCATAGCGGTCCTCATCGGAACCTGTCGAGGGTGG 124

QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysGlyGluPheGlnLeuGlnAsn 95
 Db 125 GATAATCGACCATTCAGGTGGTGTGTTGTTAGGCGGCCCTGCTGATCGCTTGGAGACC 184

QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
 Db 185 CAGCCCTTGTATCGAGGCGCTCATCGTCATGCGGCGGACGGTTATCACC 244

QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
 Db 245 ACCTTCTCCGGGCGGTGGTGTGTTGTTAGGCGGCCCTGCTGATCGCTGCTCT 304

QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu 155
 Db 305 GTCTCTTACTTTTCAGTGGCTGGTGAATGCTTTCGGGCCCAACGCTCATGTGTTGCTC 364

QY 156 LeuAlaArgPheValAspGlyPheGlyLeuAlaValThrLeuValProLeuTyr 175
Db 365 TTGGCAGGCTATCGACGGTTCGGTATCGGTTGGCTGTCACCTTGTCCCTCTTTAC 424
QY 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
Db 425 ATTTCGAGACCGCCCGACTGACATTAGAGGCTGCTGAACACGCTGCGGAGTTCACT 484
QY 196 GlySerGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
Db 485 GGCTCAGAGGATGTTCTCTTCTACTGATGGTGTTCACCATGCTGCTCATGCCGAG 544
QY 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
Db 545 CTTGACTGGAGATCAGCTTGGGTTTGTGATCCCGTGGTTATGATTTGCTATG 604
QY 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerIleGlyArgMetAlaGlu 255
Db 605 ACTGTCTTCTATTTCGCCGAGTCCCGAGATGGCTTGTGAGCAAGAGAAATGGCCGAG 664
QY 256 AlaIleGlyValLeuGlnIleLeuArgGlyAspValSerGlyLeuLeuSerLeu 275
Db 665 GCCAAGCGATGTTGCAAGACTCGCGGAGGAGAGATGTCTCAGAGAAATGGCCCT 724
QY 276 LeuLeuGluGlyLeuGluValGlyAspThrSerIleGluIleGlyPro 295
Db 725 CTTGTTGAAGGATTTGGGTTGGGAAAGACACACATTTTGAGGAATACATAATGGCCT 784
QY 296 AlaThrGluAlaAlaAspLeuValThrAspGlyAspIleGlnIleThrLeuTyr 315
Db 785 GATGATGAGCTGCTGATGACGCTCGCTGCCA---GATGAAGAGAA-GTGAAGCTGTAC 840
QY 316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerIleGlyProIleMetLeuGly 335
Db 841 GGAGCTGAAGAGGGGTA-TCTTGATCGCCCT-----CTGTTAGGG 881
QY 336 SerVal-----LeuSerLeuAlaSerArgHisGlySerMetVal 348
Db 882 CGCGTGGCAAGTGCACCTTGGAGCGTGGGTTCTCATGCTCTCATGGGAGTATGTT 941
QY 349 AsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsn 368
Db 942 AGTCAGGGTAA-TCTCTGCTGGACCCACT-GTTACTCTCTCGGAAGTGTCCATAGAGA 999
QY 369 MetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPhe 388
Db 1000 ---TCCCTGAGTAATGGGAGCAATGCGCCCAATGTTCTCTAACTT-GGCCACATGTT 1055
QY 389 Ser 389
Db 1056 ACG 1058
RESULT 10
LOCUS BQ752737
DEFINITION WHE4118_E11_I22S Wheat salt-stressed root cDNA library
aestivum cDNA clone WHE4118_E11_I22, mRNA sequence.
ACCESSION BQ752737
VERSION BQ752737.1 GI:21930519
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 829)
AUTHORS Anderson, O.D., Akhunov, E., Chao, S., Crossman, C., Deal, K.,
Dvorak, J., Lazo, G.R., Pham, J., Rausch, C.J., Wilson, C. and Woo, J.
TITLE The structure and function of the expressed portion of the wheat
genomes - Salt-stressed root cDNA library
JOURNAL Unpublished (2002)
COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific

West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595973
Fax: 5105595818
Email: candersn@pw.usda.gov
Sequences have been trimmed
quality sequence with phred score less than 20
Seq primer: SK primer.
FEATURES
Location/Qualifiers
1..829
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE4118_E11_I22"
/tissue_type="Roots"
/dev_stage="Full tillering"
/lab_host="E. coli SOLR"
/clone_lib="Wheat salt-stressed root cDNA library"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid
pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; Hydroponic
plants grown to full tillering stage were treated with 150
mM NaCl for either 12 hours or 7 days. Root tissues of the
plants subjected to both types of treatment were collected
separately at University of California, Davis (E. Akhunov
and K. Deal in J. Dvorak's lab). Total RNA was prepared
separately from the two samples (12h and 7day treatments),
and equal amount of RNA was then pooled. PolyA RNA was
purified from the pooled RNA, a cDNA library was made, and
the cDNA clones were in vivo excised to give pBluescript
SK(-) phagemids in J. Dvorak's lab (E. Akhunov, J. Dvorak)
at the University of California, Davis. Colony plating
plasmid DNA preparations and DNA sequencing were performed
in the OD Anderson lab (all other authors)."

ORIGIN

Alignment Scores:
Pred. No.: 2,18e-100 Length: 829
Score: 1101.00 Matches: 220
Percent Similarity: 89.89% Conservatives: 20
Best Local Similarity: 82.40% Mismatches: 25
Query Match: 27.02% Indels: 2
DB: 5 Gaps: 2
US-10-051-909-32 (1-800) x BQ752737 (1-829)
QY 486 GluGlyGluAsnGlyArgGlyGlyPheGlyValTyrLeuHisGlnGluGly 505
Db 28 AAAGCGAGGATGGCAAGAGGAGGCTTCAAAAGAAATCTACTTGCCACCAAGAGGG 87
QY 506 ValProGlySerArgArgGlySerIleValSerLeuProGlyGlyAspValPheGlu 525
Db 88 GTGGCGGACTCAAGAGAGGGGCTCTGTGTTTTCACCTCTGGTGGGGTGTGCCACCAA 147
QY 526 --GlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLys 544
Db 148 GGGGGCAGTGGGTTTATACACGCTGCTGTTTGGTAAGCCACTCGGCTCTTTACTCAAAG 207
QY 545 GlyLeuAlaGluProArgMet---SerAspAlaAlaMetValHisProSerGluValAla 563
Db 208 GATCTTATGAAAGAGCGTATGGCGCGCTGATGCGCGCATCATCCATTCATGGAGGAGCT 267
QY 564 AlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuVal 583
Db 268 CCCAAGGTTCAATCTGGAAGATCTGTTGACCTGGTGTGAGGGGTGCTATTTGCTGTC 327
QY 584 GlyValGlyIleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThr 603
Db 328 GGGGTTGGAATTCAGATGCTTCAGCAGTGTGCTGGAATAAATGAGTCTCTTACTATATCT 387
QY 604 ProGlnIleGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSer 623
Db 388 CCTCAAAATCTTGAGCAAGCTGTGGCGGTTCTTCTTCTTCCAACTTGGCCTCAGTTCA 447

QY 624 AlaSerAlaSerIleLeuLeuSerSerIleThrThrLeuLeuMetLeuProCysIleGly 643
 Db 448 GCATCAGCGTCCATCTTGATCAGCTCTCTCCACCTTACTCATGCTCCCAAGCATTTGGT 507
 QY 644 PheAlaMetLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIlePro 663
 Db 508 GTAGCCATGAGACTTATGGATATATCTGGAAGAGGTTTCTTCTACTGGGCACATCCC 567
 QY 664 IleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeu 683
 Db 568 ATCTGTGATAGCATCCCTAAATTTGTTGGTTGTGTCCCAATGTTTCAACTTGAGTACGGTG 627
 QY 684 AlaHisAlaLeuLeuSerThrValSerValIleValThrPheCysCysPheValMetGly 703
 Db 628 CCCACGCTGTCTCTCCACAGTTAGGTCATGTTCTACTTCTGCTCTTGTTCATGGGC 687
 QY 704 PheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeu 723
 Db 688 TTTGGCCCAATCCCAACATCTATGTCGACAGATTTTCCCCACAGAGTCCCGTGTGTC 747
 QY 724 CysIleAlaIleCysAlaPheThrPheThrIleGlyAspIleIleValThrThrSerLeu 743
 Db 748 TGCATCGCTATTTCGGCCCTCACATCTGGATATGTGACATATTGTTACCTACAGCCTG 807
 QY 744 ProValMetLeuAsnAlaIle 750
 Db 808 CCTGTGATGCTGAATGCTATT 828

RESULT 11

CA170109 871 bp mRNA linear EST 24-SBP-2003
 LOCUS SCQGSB1083B09.9 SB1 Saccharum officinarum cDNA clone SCQGSB1083B09
 DEFINITION 5', mRNA sequence.

ACCESSION CA170109

VERSION CA170109.1 GI:35093205

KEYWORDS EST.

SOURCE Saccharum officinarum

ORGANISM Saccharum officinarum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
 complex.

1 (bases 1 to 871)

Vettore,A.L.; da Silva,F.R.; Kemper,E.L. and Arruda,P.

The libraries that made SUCEST

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: parruda@unicamp.br

Clone distribution: clone distribution information can be found

through the Brazilian Clone Collection Center (BCCC) at

http://www.bcccenter.fcav.unesp.br

Plate: 683 row: B column: 09

Seq primer: T7 Promoter Primer.

Location/Qualifiers

1..871

/organism="Saccharum officinarum"

/mol_type="mRNA"

/db_xref="taxon:4547"

/clone="SCQGSB1083B09"

/lab_host="DH10B"

/clone_lib="SB1"

/notes="Organ: Stalk Bark from adult plants; Vector:

pSport1; Site 1: SalI; Site 2: NotI; An unidirectional

cDNA library generated from [stalk Bark from adult

plants]. cDNA was prepared from polyA+ mRNA using

Superscript Plasmid System Kit (Invitrogen). The

double-strand cDNAs were fractionated in a sepharose

Cl-2B 40cm-columns and fragments sizing between 0.8 and
 1.5 Kb were directionally cloned into the vector. Details
 of each source of RNA and library construction can be
 obtained at <http://sucest.lad.ic.unicamp.br/public>

ORIGIN

Alignment Scores:
 Pred. No.: 2,98e-100 Length: 871
 Score: 1100.00 Matches: 215
 Percent Similarity: 86.11% Conservative: 33
 Best Local Similarity: 74.65% Mismatches: 40
 Query Match: 26.99% Indels: 2
 DB: 6 Gaps: 0

US-10-051-909-32 (1-800) x CA170109 (1-871)

QY 420 AsptYrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGly 439
 Db 9 GACTATGAAACAATGATACAGAGCCCTCTGCTGTCCCGACAGAAATACAAGTCCCAAGG 68
 QY 440 LysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgGlnThrLeu 459
 Db 69 AGGGCATTCACATCACGGCACCGTGGAAAGTCTTTGGTTTGAGAAGAAGAGCTG 128
 QY 460 LeuGlyGluGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAla 479
 Db 129 TCCGATGAGGTTGGTGGGAGCAACAGCACTGGCATTTGGTGGCGATGGCAGCTGCC 188
 QY 480 TrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgVal 499
 Db 189 TGGAAATGCTCAGAGCCGAGAGGTGAGACGGTAAGAGGAAGGAGGTTCAAAGAATC 248
 QY 500 TyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGly 519
 Db 249 TACTTACACCAAGAGGAGTTGCTGGCTCAAGAGGGGATCTGTTGCTCCTCCTCGT 308
 QY 520 GlyGlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSer 539
 Db 309 CGAGGTGATGCTCTGAAGGTGTAAGTTTATATATCTGCTGCTGCTGCTGCTGCTGCT 368
 QY 540 AlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisPro 559
 Db 369 CCTCTTACCAGGGATATTACAGAACAGCGCATGCCCGTCCAGCTACAATGCCACCA 428
 QY 560 SerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArg 579
 Db 429 TCAGAGGAGCTGCCAAAGTTTCCAAAGCTGGAGAGACCTGTTTGAACCTGCTGTGAGAC 488
 QY 580 AlaLeuLeuValGlyValGlyIleGlnIleGlnGlnPheAlaGlyIleAsnGlyVal 599
 Db 489 CCCTGTTGTTGTAATGGAATTCAGATCCTTCAACAGTTCGGGGAATAAATGNGTT 548
 QY 600 LeuTyrTrpThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPhe 619
 Db 549 CTCTACTACATCCGCAAAATTATGGAGCAAGCTGGTGGCAGTCTTATTTCCTCAATCTT 608
 QY 620 GlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeu 639
 Db 609 GGTCTCAGTTCCGATCAGATCCATCCTAATTAAGTTCTGTCTCCTGCTGCTGCTGCTG 668
 QY 640 ProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeu 659
 Db 669 CCTAGCATTTGGTTAGCCATGAGACTTATGAGCATCTATGGAAGAGGTTTCTGCTGCTA 728
 QY 660 GlyThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAsp 679
 Db 729 AGCAAAATCCAGTCTTGATAGCATCTTTAATTTGCTGCTGCTGCTGCTGCTGCTGCTG 788
 QY 680 LeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCys 699
 Db 789 TTGGGTACAGTGGTCCATGCTGTGCTCTCCACATT-AGTGTCTACCTACTTATGCTGTC 847
 QY 700 PheValMetGlyPheGlyProIle 707

[illegible]

RESULT 13
CF445664
LOCUS
848 bp mRNA linear EST 04-SEP-2003
DEFINITION EST62009 normalized cDNA library of onion Allium cepa cDNA clone
ACCESSION
CF445664
VERSION
CF445664.1 GI:34468366
KEYWORDS
EST.
SOURCE Allium cepa (onion)
ORGANISM Allium cepa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;
Allium.

```

REFERENCE
1 (bases 1 to 848)
AUTHORS
Havey,M.J., Cheung,F., Van Aken,S., Utterback,T. and Town,C.D.
TITLE
Expressed Sequence Tags from a normalized library of mixed onion
tissues (Allium cepa)
JOURNAL
Unpublished (2003)
COMMENT
Contact: Havey MJ
Department of Horticulture
USDA-ARS and University of Wisconsin
1575 Linden Drive, Madison, WI 53706, USA
Tel: 608-262-1830
Fax: 608-262-4743
Email: mjhavey@facstaff.wisc.edu
TIGR sequence name ACAL196TR. For more information:
http://haveylab hort.wisc.edu
Seq primer: CAG GAA ACA GCT ATG ACC.
FEATURES
Location/Qualifiers
1..848
/organism="Allium cepa"
/mol_type="mRNA"
/cultivar="Red Creole (bulbs), unknown(callus), Ebano &
Texas Legend(roofs)"
/db_xref="taxon:4679"
/clone="ACAL196"
/tissue_type="Callus, roots, and young bulbs"
/clone_lib="normalized cDNA library of onion"
/notes="Vector: pCMVSPORT6.1-ccdb (Invitrogen); Site 1:
EcoRV (5'); Site 2: NotI (3'); Equal molar amounts of mRNA
from callus, roots, and young bulbs were combined to
synthesize the library. Normalization to enrich for
low-copy transcripts was performed by proprietary
techniques of Invitrogen."
ORIGIN
Alignment Scores:
Pred. No.: 7,04e-97 Length: 848
Score: 1066.50 Matches: 215
Percent Similarity: 82.53% Conservative: 26
Best Local Similarity: 73.63% Mismatches: 36
Query Match: 26.17% Indels: 15
DB: 7 Gaps: 4
US-10-051-909-32 (1-800) x CF445664 (1-848)
QY 477 GlnLeuAlaTrpLysTrpSerGluLysGluGlyGluLysGluGlyGlyPhe 496
DB 1 CAGCTTGTCTGCGAAAGAGTGAAGGGGAGGACAGAAATGGTAAAGGAAGTGGGTTT 60
QY 497 LysArgValTrpLeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSer 516
DB 61 AAAAGGATATATTTACAT---GAGGGGTGCTGCTGTTCTAGAGAGGTCATTTGTTCC 117
QY 517 LeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaLeuVal 536
DB 118 ATACAG---GGAGGTGATGTTCCGGAAGATAGCAGTACGTGCGAGGTGGCAGCTTTGGTT 174
QY 537 SerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMet 556
DB 175 AGTGAGTCTGCTCTTTATTCGAAG-----GATAGAGTGGG 210
QY 557 ValHisPro-----SerGluValAlaAlaLysGlySerArgTrpLysAsp 571
DB 211 CATCATCTGTGGCTATGTTGATTCTAATACGCGAGTAAGGCAAGCAGATGTCGAT 270
QY 572 LeuPheGluProGlyValArgAlaLeuValGlyValGlyIleGlnIleLeuGln 591
DB 271 CTTCTTGAACACAGGTGTTAGAGGTCTTTGATTGTTGGAGTTGGATTCAGATTCCTCAG 330
QY 592 GlnPheAlaGlyIleAsnGlyValLeuTrpTrpProGlnIleLeuGluGlnAlaGly 611
DB 331 CAGTTTTCAGGCATCAACCGGTGTTCTGTACTACACACACAGATTCCTTGACACAGCAGGT 390
QY 612 ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerIleLeuIleSer 631

```

```

DB 391 GTTGAGTTCTCCTCTCAAACTCGGACTCACTGATTCTTCAGCTCCACTCTTTATAAGT 450
QY 632 SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu 651
DB 451 GCATTGGTTACTCTACTAATGCTTCATCTCGGTATTCGTATGAGATTTCATGACATA 510
QY 652 SerGlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671
DB 511 GCTGGAGAGAGGTTCTTCTTCTATCCCAATACCCATCTTAATCTTATCTTTAGTCATA 570
QY 672 LeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691
DB 571 CTGTAATATCCACATATTCAGATTCGGACAGTGGTACATGCTATCTTTTCAACTGCA 630
QY 692 SerValIleValTrpPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu 711
DB 631 AGTGATGATGATATCTCTGTTGTTTGTAGGGTTTGGACCATACCCACATCTTA 690
QY 712 CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731
DB 691 TGTGCTGTAATATTCCTCAACAGAGTCAGAGGCATCTGCTATCTGCTGCTCTCAG 750
QY 732 PheTrpIleGlyAspIleValThrTrpSerLeuProValMetLeuAsnAlaIleGly 751
DB 751 TTCTGGATCGGGGACATATATGTCATATACCTTCTCTGATGCTCACTCCATTGGT 810
QY 752 LeuAlaGlyValPheSerIleTrpAlaValValCys 763
DB 811 CTGGCTGGGGTTTGTGTTATTTACGCTACTGCTGT 846
RESULT 14
CA243478
LOCUS
DEFINITION
5', mRNA sequence.
ACCESSION
CA243478
VERSION
CA243478.1 GI:35320137
KEYWORDS
EST.
SOURCE
Saccharum officinarum
ORGANISM
Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
REFERENCE
1 (bases 1 to 769)
Vettore,A.L., da Silva,F.R., Kemper,B.L. and Arruda,P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bccccenter.fcav.unesp.br
Plate: 071 row: E column: 03
Seq primer: T7 Promoter Primer.
FEATURES
Location/Qualifiers
1..769
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCQFL4071E03"
/lab_host="DH10B"
/clone_lib="FL4"
/notes="Organ: developed inflorescence and rachis
(20cm-long); Vector: pSport1; Site 1: SalI; Site 2: NotI;
An unidirectional cDNA library generated from [developed
inflorescence and rachis (20cm-long)]. cDNA was prepared
from polyA+ mRNA using SuperScript System Kit

```



```
QY 352 ValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGln 371
Db |||||
Db 241 GTACCCCTTATGGATCCGATTGTGACACTTTTGGGAGTGTCCACGAGAATATGCTCAA 300
QY 372 AlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThr 391
Db |||||
Db 301 GCTGGAGCAAGTATGAGGAGCACATTGTTCCAAACTTTGGAAGTATGTTCAAGTGTCACA 360
QY 392 AspGlnHisAlaLysAsnGluGlnTrpAspGluLysLeuHisArgAspAspGluGlu 411
Db |||||
Db 361 GATCAGCATGCCAAAAATGACAGTGGGACGAGAGAATCTTCACAGGGACGATGAGGAG 420
QY 412 TyrAlaSerAspGlyAlaGlyGlyValAspTyrGluAspAsnLeuHisSerProLeuLeuSer 431
Db |||||
Db 421 TATGCACTGTGATGGTGCAGGAGGTGATTATGAGGACATCTCCACAGCCCATTTGCTGTC 480
QY 432 ArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAla 451
Db |||||
Db 481 AGGCAGACACACAGTGTGGAAGGAAAGACATTGTGCACCATTGTCACCGTGGAGTGCT 540
QY 452 LeuSerMetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAsp 471
Db |||||
Db 541 TTGAGCATGAGAGGCAAGCCTCTTNGGGAGGCTTGAAGGGTGTGAGCAGCAGCATGAT 600
QY 472 IleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArg 491
Db |||||
Db 601 ATTGGTGGAGGATGGCAGCTTGCATGCAATGGCAN-----AGANGGAAGTGAGGATGG 654
QY 492 LysGluGlyGlyPheLysArgValTyrLeuHisGlnGlu 504
Db |||||
Db 655 TAAGAAGGAGTGGTTCAAAGAGTTCTTACACCAAGAG 693
```

Search completed: October 13, 2004, 19:46:27
Job time : 6526.01 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 11:33:18 ; Search time 89.8744 Seconds
(without alignments)
5121.591 Million cell updates/sec

Title: US-10-051-909-32

Perfect score: 4075

Sequence: 1 IRSGSLAVQTPTPLDRR.....PLEVITEFAVGAQAQA 800

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 02.1*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3397	83.4	746	2	O6H536
2	3153	77.4	745	2	O7X9V6
3	2834	69.5	740	2	O7XCN9
4	2834	69.5	740	2	O9FRK0
5	2808	68.9	743	2	O8GT52
6	2654	65.1	739	2	O8LP08
7	2584	63.4	729	2	O65497
8	2278	55.9	734	2	O9SY03
9	2275	55.8	734	2	O9LM67
10	2245	55.1	723	2	O9SD00
11	2191.5	53.8	753	2	O8GT51
12	2183	53.6	734	2	O96290
13	1963	48.2	652	2	O6K967
14	1763	43.3	714	2	O8H887
15	1763	43.3	723	2	O9SNK7
16	1270.5	31.2	745	2	O8LJX7
17	1050	25.8	643	2	O8H6J2
18	621.5	15.3	664	2	O9XEV1
19	587.5	14.4	457	2	P96742
20	542	13.3	506	2	O7XU54
21	536	13.2	580	2	O9C757
22	535.5	13.1	581	2	O9LKH2
23	533	13.1	461	1	C9BC BACSU
24	527.5	12.9	498	2	O84U74
25	524.5	12.9	580	2	O9ZQ06
26	520.5	12.8	582	2	O23492
27	519	12.7	581	2	O9LKH1
28	518.5	12.7	469	2	O9LJH7
29	512.5	12.6	577	2	O945E5
30	512	12.6	511	2	O6ZFM6
31	512	12.6	511	2	BAC83311

32	510.5	12.5	519	2	O7XB36
33	505	12.4	473	2	O34718
34	505	12.4	510	2	O6ZFM7
35	505	12.4	510	2	BAC83310
36	501.5	12.3	468	2	O8AIQ3
37	500	12.3	509	2	O8VZR6
38	498	12.2	463	2	O8EMPS
39	494	12.1	470	2	O9LLE0
40	493.5	12.1	508	2	O9ZNS0
41	493	12.1	493	2	O23213
42	492	12.1	467	2	O8COA7
43	492	12.1	534	2	O9LLE1
44	491.5	12.1	468	2	O8AAM8
45	490.5	12.0	538	2	O84QH3

ALIGNMENTS

RESULT 1

Q6H536	PRELIMINARY;	PRT;	746 AA.
ID	O6H536		
AC	O6H536;		
DT	05-JUL-2004 (Tremblrel. 27, Created)		
DT	05-JUL-2004 (Tremblrel. 27, Last sequence update)		
DT	05-JUL-2004 (Tremblrel. 27, Last annotation update)		
DE	Putative hexose transporter.		
GN	Name=OSUNB0035N08.17;		
OS	Oryza sativa (japonica cultivar-group).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
OC	Eriharoidae; Oryzae; Oryza.		
OX	NCBI_TaxID=39947;		
RP	SEQUENCE FROM N.A.		
RA	Sasaki T., Matsumoto T., Katayose Y.;		
RL	Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.		
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).		
CC	-!- SIMILARITY: Belongs to the sugar transporter family.		
DR	EMBL; AF005756; BAD26163.1; -		
DR	InterPro; IPR007114; MFS.		
DR	InterPro; IPR005086; Peptidase S26.		
DR	InterPro; IPR005828; Sub transporter.		
DR	InterPro; IPR003663; Sugar transp.		
DR	InterPro; IPR005829; Sug transporter.		
DR	Refam; PF00083; Sugar tr_1		
DR	PRINTS; PRO0171; SUGTRNSPORT.		
DR	PROSITE; PS00850; MFS; 1.		
DR	PROSITE; PS00501; SPASE I_1; UNKNOWN 1.		
DR	PROSITE; PS00216; SUGAR TRANSPORT 1; 1.		
DR	PROSITE; PS00217; SUGAR TRANSPORT 2; 1.		
KW	Sugar transport; Transmembrane; Transport.		
SQ	SEQUENCE 746 AA; 80108 MW; 61F5890E2BFED02E CRC64;		

Query Match 83.4%; Score 3397; DB 2; Length 746;
Best Local Similarity 88.2%; Pred. No. 2.5e-212;
Matches 655; Conservative 40; Mismatches 48; Indels 0; Gaps 0;

QY	56	MSGAVLVAIVASIGNLQGNNTAAAVLYIKKEFQLQNEPTVEGLVMSLIGATIVT	115
DB	1	MSGAAVLAIVASIGNLQGNNTAAAGVLYIKKEFQLQNEPTVEGLVMSLIGATIT	60
QY	116	TFSGPLSDIGRRPMLILSSILYFFSGHLMSPNNVYLLARFVDFGFGIGLAVTLVPL	175
DB	61	TFSGPVSVDIGRRPMLILSSILYFLSSLLMLSPNNVYLLARLIDFGIGLAVTLVPL	120
QY	176	ISEIAPSEIRGLNLTLPQFSGSGGNFLSCVFGNSLSPSPDWRIMLGVAIPSPFGL	235
DB	121	ISEIAPSEIRGLNLTLPQFSGSGGNFLSCVFGNSLSPSPDWRIMLGVAIPSPFGL	180
QY	236	TIFFYLPESPRMLVSKGRMAEAKVILQKLRGKDDVSGELSLLEGLGVGDTSIEEYIIGP	295
DB	181	TIFFYLPESPRMLVSKGRMAEAKVILQKLRGKDDVSGELSLLEGLGVGDTSIEEYIIGP	240

296 ATEAADLVTGDKKEQITLYCPGEGQSWIARPSKGPIMLGSLVSLASRRHSGSMVNSQVPLM 355
 Db QY
 241 AIEPADRHVVDGDKQITLYCPGEGQSWIARPSKGPISILGSLVSLASRRHSGSMVNSQVPLM 300
 Db QY
 356 DPIVTLFGSVHNPQAGGSMRSTLFPNFGSMFVTDQHKAKNQWDEENLHRDDEYASD 415
 Db QY
 301 DPIVTLFGSVHNPQAGGSMRSTLFPNFGSMFVTDQHKPKVDQWDEENLHRDDEYASD 360
 Db QY
 416 GAGDYEDNLHSPILSRQAATGAEKQITVHHGHGSGSALSMSRQTLGCGGSDGVSSDTIGGG 475
 Db QY
 361 GAGDYEDNLHSPILSRQTTSAEKGQIAHAGHSGSALSMSRRLLEGGGSAVSTGIGG 420
 Db QY
 476 WQIAWKSEKEGNGRKEGKFKRVLHQBVPQSGRSGSVLSLPGGSDVFGSGFVHAAL 535
 Db QY
 421 WQIAWKSEKEGNGRKEGKFKRIYHQBVPQSGRSGSVLSLPGGSDVFGSGFVHAAL 480
 Db QY
 536 VSQALPSKGLABPRMSDAAMVHPSEVAAGKSEWKDLFEPGVRRALLVGVGIIQIQFAG 595
 Db QY
 481 VSQALPSKGLABPRMSDAAMVHPSEVAAGKSEWKDLFEPGVRRALLVGVGIIQIQFAG 540
 Db QY
 596 INGVLVYTPQILEQAGVAVILSKFGSSASASILISLTLTLLMPCIGFAMLLMDLSGRR 655
 Db QY
 541 INGVLVYTPQILEQAGVAVILSKFGSSASASILISLTLTLLMPCIGFAMLLMDLSGRR 600
 Db QY
 656 FLIGTTPILIASIVILVNSLIDLGTALHALSTVSVIVYFCCFVWGFGPIPNILCAEI 715
 Db QY
 601 FLIGTTPILIASIVILVNSLIDLGTALHALSTVSVIVYFCCFVWGFGPIPNILCAEI 560
 Db QY
 716 FPTRVRGLCIACAFTWIGDIIIVTSLPVLNAIGLAGVFSIYAVVCSISFVFLKVP 775
 Db QY
 661 FPTRVRGLCIACAFTWIGDIIIVTSLPVLNAIGLAGVFSIYAVVCSISFVFLKVP 720
 Db QY
 776 ETGKMPLEVITEFFAVGAKQAAA 798
 Db QY
 721 ETGKMPLEVITEFFAVGAKQAAA 743
 Db QY

RESULT 2

Q7X9Y6 PRELIMINARY; PRT; 745 AA.
 ID Q7X9Y6
 AC Q7X9Y6
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative sugar transporter type 2a.
 OS Saccharum hybrid cultivar.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC PACCAD clade; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Saccharum officinarum complex.
 OC NCBI_TaxID=128810;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22740598; PubMed=12856943;
 RA Casu R.E., Grof C.P.L., Rae A.L., McIntyre C.L., Dimmock C.M.,
 RA Manners J.M.;
 RT "Identification of a novel sugar transporter homologue strongly
 RT expressed in maturing stem vascular tissues of sugarcane by expressed
 RT sequence tag and microarray analysis."
 RL Plant Mol. Biol. 52:371-386(2003).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 DR EMBL; AY165599; A037640.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005351; F:sugar porter activity; IEA.
 DR GO; GO:0005315; F:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub_transporter.
 DR InterPro; IPR005829; Sug_transporter.
 DR Pfam; PF00083; Sugar_tr; 1.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.

KW Sugar transport; Transmembrane.
 SQ SEQUENCE 745 AA; 79955 MW; 54B12A7BFBDB8B155 CRC64;
 Query Match 77.4%; Score 3153; DB 2; Length 745;
 Best Local Similarity 80.7%; Pred. No. 1.9e-196;
 Matches 601; Conservative 67; Mismatches 77; Indels 0; Gaps 0;
 QY 56 MSGAVLVAIVASIGNLQGWDNATIAAALVYIKKEFQIQNEPTVEGLIVMSLIGATIT 115
 Db 1 MSGAALVAIVASIGNLQGWDNATISAAALYIKKEFLESEPTVEGLIVMSLIGATIT 60
 QY 116 TFSGPLSDSICRRPMLILSSILYFFSGILMLWSNVYVLLARFVDGFGILAVTLVPLV 175
 Db 61 TFSGPVSDWIGRLPMLILSSVLYFVSSILMLWSNVYVLLARLNGFGVGLAVTLVPLV 120
 QY 176 ISEIAPSIIRGLANTLPQFGSGGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPLSPFFGL 235
 Db 121 ISETAPPRIIRGLANTLPQFGSGGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPLSPFFGL 180
 QY 236 TIFVLPSPRLVSKGRMAEAKVLOKRGKDDYSGELSLLEGLVGGDTSEIYIIGP 295
 Db 181 TIFVLPSPRLVSKGRMAEAKVLOKRGKDDYSGELSLLEGLVGGDTSEIYIIGP 240
 QY 296 ATEAADLVTGDKKEQITLYCPGEGQSWIARPSKGPIMLGSLVSLASRRHSGSMVNSQVPLM 355
 Db 241 AIEPADRHVVDGDKQITLYCPGEGQSWIARPSKGPISILGSLVSLASRRHSGSMVNSQVPLM 300
 QY 356 DPIVTLFGSVHNPQAGGSMRSTLFPNFGSMFVTDQHKAKNQWDEENLHRDDEYASD 415
 Db 301 DPIVTLFGSVHNPQAGGSMRSTLFPNFGSMFVTDQHKPKVDQWDEENLHRDDEYASD 360
 QY 416 GAGDYEDNLHSPILSRQAATGAEKQITVHHGHGSGSALSMSRQTLGCGGSDGVSSDTIGGG 475
 Db 361 GAGDYEDNLHSPILSRQTTSAEKGQIAHAGHSGSALSMSRRLLEGGGSAVSTGIGG 420
 QY 476 WQIAWKSEKEGNGRKEGKFKRVLHQBVPQSGRSGSVLSLPGGSDVFGSGFVHAAL 535
 Db 421 WQIAWKSEKEGNGRKEGKFKRIYHQBVPQSGRSGSVLSLPGGSDVFGSGFVHAAL 480
 QY 536 VSQALPSKGLABPRMSDAAMVHPSEVAAGKSEWKDLFEPGVRRALLVGVGIIQIQFAG 595
 Db 481 VSQALPSKGLABPRMSDAAMVHPSEVAAGKSEWKDLFEPGVRRALLVGVGIIQIQFAG 540
 QY 596 INGVLVYTPQILEQAGVAVILSKFGSSASASILISLTLTLLMPCIGFAMLLMDLSGRR 655
 Db 541 INGVLVYTPQILEQAGVAVILSKFGSSASASILISLTLTLLMPCIGFAMLLMDLSGRR 600
 QY 656 FLIGTTPILIASIVILVNSLIDLGTALHALSTVSVIVYFCCFVWGFGPIPNILCAEI 715
 Db 601 FLIGTTPILIASIVILVNSLIDLGTALHALSTVSVIVYFCCFVWGFGPIPNILCAEI 560
 QY 716 FPTRVRGLCIACAFTWIGDIIIVTSLPVLNAIGLAGVFSIYAVVCSISFVFLKVP 775
 Db 661 FPTRVRGLCIACAFTWIGDIIIVTSLPVLNAIGLAGVFSIYAVVCSISFVFLKVP 720
 QY 776 ETGKMPLEVITEFFAVGAKQAAA 800
 Db 721 ETGKMPLEVITEFFAVGAKQAAA 745

RESULT 3

Q7XCM9 PRELIMINARY; PRT; 740 AA.
 ID Q7XCM9
 AC Q7XCM9
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative sugar transporter.
 DE ORFNames=OSJNEB0064P21.3;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoidae; Oryzae; Oryza.

OX NCBI_TaxID=39947;
RN [1]
RA SEQUENCE FROM N.A.
RA The Rice Chromosome 10 Sequencing Consortium;
RT "in-depth view of structure, activity, and evolution of rice
RT chromosome 10";
RL Science 300:1566-1569(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR EMBL; AE017116; AAP54842.1; -
DR Gramene; Q7XCM9; -
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0005351; P: sugar porter activity; IEA.
DR GO; GO:0005215; P: transporter activity; IEA.
DR GO; GO:0006810; P: transporter; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR InterPro; IPR005829; Sug transporter.
DR Pfam; PF00083; Sugar tr; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane.
SQ SEQUENCE 740 AA; 78898 MW; D4C003E0A4980C46 CRC64;

Query Match 69.5%; Score 2834; DB 2; Length 740;
Best Local Similarity 74.5%; Pred. No. 1e-175;
Matches 553; Conservative 75; Mismatches 108; Indels 6; Gaps 5;

QY 56 MSGAVLVAIVASIGNLQGDNDNATIAAAVLYIKKEFQLENEPTVEGLIVMSLIGATIVT 115
DB 1 MAGAVLVAIAASIGNLQGDNDNATIAAGAVLYIKKEFNQSEPLIEGLIVAMSLIGATIT 60

QY 116 TFSGPLSDSISGRPRMLILSSILYFSGILMSPNVYVILLARFVDGEGIGLAVTLVPLY 175
DB 61 TFSGAVDSFGRPRMLIASAVLYFVSGVLMWAPNVYVILLARLDGGGIGLAVTLVPLY 120

QY 176 ISEIAPSEIRGLNLTLPQFSGSGGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPLSFFFLG 235
DB 121 ISETAPTDIRGLNLTLPQFSGSGGMFLSYCMVFGMSLMPQPDWRIMLGVLSPISLIYPAL 180

DB 596 RELLATIPILVALAILLVNILDVGMVHASTVSVILYFCFFVNGFGPIPNILCAE 655
QY 715 IFFTVRGLCTAICAFTHWIGDIIYTSYSLPWNMINAGLAVESIVAVVCLISFVFEVLKV 774
DB 656 IFFTVRGICIAICAFTHWIGDIIYTSYSLPWNMINAGLAVESIVAVVCLISFVFEVLKV 715
QY 775 PETKGMPLVETEPFPAVGAQKA 796
DB 716 PETKGMPLVETEPFPAVGAQKA 737

RESULT 4
QYFRKO PRELIMINARY; PRT; 740 AA.
AC QYFRKO;
DT 01-VAR-2001 (TREMELrel. 16, Created)
DT 01-VAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Putative sugar transporter.
GN Name=OSJNB0064P21.3;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Burr P.C., Hsiao J.,
RA Zismann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E.,
RA Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,
RA Quackenbush J., White G., Salzberg S.L., Fraser C.M.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL; AC073166; AAG46115.1; -
DR Gramene; QYFRKO; -
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0005351; P: sugar porter activity; IEA.
DR GO; GO:0005215; P: transporter activity; IEA.
DR GO; GO:0008643; P: carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR InterPro; IPR003663; Sugar transpt.
DR InterPro; IPR005829; Sug transporter.
DR Pfam; PF00083; Sugar tr; 1.
DR PRINTS; PR00171; SUGTRANSPORT.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 740 AA; 78898 MW; D4C003E0A4980C46 CRC64;

Query Match 69.5%; Score 2834; DB 2; Length 740;
Best Local Similarity 74.5%; Pred. No. 1e-175;
Matches 553; Conservative 75; Mismatches 108; Indels 6; Gaps 5;

QY 56 MSGAVLVAIVASIGNLQGDNDNATIAAAVLYIKKEFQLENEPTVEGLIVMSLIGATIVT 115
DB 1 MAGAVLVAIAASIGNLQGDNDNATIAAGAVLYIKKEFNQSEPLIEGLIVAMSLIGATIT 60

QY 116 TFSGPLSDSISGRPRMLILSSILYFSGILMSPNVYVILLARFVDGEGIGLAVTLVPLY 175
DB 61 TFSGAVDSFGRPRMLIASAVLYFVSGVLMWAPNVYVILLARLDGGGIGLAVTLVPLY 120

QY 176 ISEIAPSEIRGLNLTLPQFSGSGGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPLSFFFLG 235
DB 121 ISETAPTDIRGLNLTLPQFSGSGGMFLSYCMVFGMSLMPQPDWRIMLGVLSPISLIYPAL 180

QY 236 TIFYPSPRWLVSKGRMAEAKVYQKLRGKDDVSGELSLLEGLVGGDTSEIYIIGP 295
D 181 TIFYPSPRWLVSKGRMAEAKVYQKLRGKDDVSGELSLLEGLVGGDTSEIYIIGP 240
QY 256 ATEAADLVTDDKEQITLYCPREGSQWAPSPKPIWGLSVLSASPHGSMVNSQVPLM 355
D 241 DDELADEGLAP-DEKIKLYGPEGLGKAPVHQQSALGALGLISRHGSMVNSQVPLM 299
QY 356 DPVITLFGSVHNNPQAGGSMRSTLFFNFGSMFVTDQAHKNEQWDEENLHRDEEYASD 415
D 300 DPVITLFGSVHNNPQAGGSMRSTLFFNFGSMFVTDQAHKNEQWDEENLHRDEEYASD 358
QY 416 GAGDYEDNLHSPILSRQATGAEGKDI-VHGHGSGSALSMRRTLLGEGGDSSTDIGG 474
D 359 HGGDDIEDSLQSPILSRQATGAEGKDI-VHGHGSGSALSMRRTLLGEGGDSSTDIGG 416
QY 475 GWQLAWKWESEGENGRKEGFKRVLHQRGVPGSRGSIIVSLPGGDDVPEGSEFVHAAA 534
D 417 GWQLAWKWESEGENGRKEGFKRVLHQRGVPGSRGSIIVSLPGGDDVPEGSEFVHAAA 475
QY 535 LVSSALFSGKLAPRMSDAMVHPSEVAAKGRWKDLFEPGVRRALLVGVGIILOQPA 594
D 476 LVSSALFSGKLAPRMSDAMVHPSEVAAKGRWKDLFEPGVRRALLVGVGIILOQPA 535
QY 595 GINGVLYTTPQILBOAGVAVILSKFGLSSASASILISLTLMLPCIGFAMLLMDLSGR 654
D 536 GINGVLYTTPQILBOAGVAVILSKFGLSSASASILISLTLMLPCIGFAMLLMDLSGR 595
QY 655 RFLLLGTIPILIASLVLVNSLIDLTALHALLSTVSVIYCCFVWFGPIPNILCAE 714
D 596 RFLLLGTIPILIASLVLVNSLIDLTALHALLSTVSVIYCCFVWFGPIPNILCAE 655
QY 715 IFPRVGLCTAICAFWIGDIIIVTSLPMLNAIGLAGVSTIYAVVCLISFVFLKV 774
D 556 IFPRVGLCTAICAFWIGDIIIVTSLPMLNAIGLAGVSTIYAVVCLISFVFLKV 715
QY 775 PETKGMPLVITEFFAVGAKQA 796
D 716 PETKGMPLVITEFFAVGAKQA 737

RESULT 5

Q8GT52 ID Q8GT52 PRELIMINARY; PRT; 743 AA.
AC Q8GT52;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hexose transporter.
GN Name=stpl;
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Developing caryopsis;
RX MEDLINE=22424051; PubMed=12535352;
RA Weschke W., Panitz R., Gubatz S., Wang Q., Radchuk R., Weber H.,
RA Wobus U.;
RT "the role of invertases and hexose transporters in controlling sugar
RT ratios in maternal and filial tissues of barley caryopses during early
RT development";
RL Plant J. 33:395-411(2003).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL: AJ534445; CAD58958.1;-;
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0005351; F:sugar porter activity; IEA.
DR GO: GO:0005215; F:transporter activity; IEA.
DR GO: GO:0008643; P:carbohydrate transport; IEA.
DR InterPro: IPR007114; NPS.

DR InterPro: IPR005828; Sub_transporter.
DR InterPro: IPR003663; Sugar_transpt.
DR InterPro: IPR005829; Sug_transporter.
DR Pfam: PF00083; Sugar_tr.1.
DR PRINTS: PR00171; SUGETRNSPT.
DR PROSITE: PS00850; MRS; 1.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transporter; Transmembrane; Transport.
SQ SEQUENCE 743 AA; 79339 MW; 0169B6BD06007DF5 CRC64;
Query Match 68.9%; Score 2808; DB 2; Length 743;
Best Local Similarity 73.1%; Pred No. 5 2e-174; Indels 12; Gaps 6;
Matches 543; Conservative 84; Mismatches 104;
QY 57 SGAVLVAIVASIGNLLQGDNATIAAAYLYIKKQFQQLQNEFTVEGLIVMSLIGATIVTT 116
D 3 SGAVLVAIVASIGNLLQGDNATIAAAYLYIKKQFQQLQNEFTVEGLIVMSLIGATIVTT 62
QY 117 FSGPLSDSISGRPMILILSSILYFFSGLTLMASPNVYVLLARFVDGFGIGLAVTLVLYI 176
D 63 FSGAVADAVGRRLTIASSVLYFVGLVNLWAPNVYVLLARLIDGFGIGLAVTLVLYI 122
QY 177 BIIAPSEIRGLINTLPQFSGSGGMFLSCMVFGMSLSPDWRIMLGVIATIPSLFFGOLT 236
D 123 SETAFTDIRGLINTLPQFSGSGGMFLSCMVFTWSLMPQDWRIMLGVLSPSLMYPALT 182
QY 237 FYLPESPRWLVSQGRMAEAKVYQKLRGKDDVSGELSLLEGLVGGDTSEIYIIGPA 296
D 183 FYLPESPRWLVSQGRMAEAKVYQKLRGKDDVSGELSLLEGLVGGDTSEIYIIGPD 242
QY 297 TEAADLVTDDKEQITLYCPREGSQWAPSPKPIWGLSVLSASPHGSMVNSQVPL 354
D 243 DELADDGLAP-DOEKLYGAEVGSMTARPVKGGQSGALGALGLMSRHGSMVNSQVPL 301
QY 355 MDPIVTLFGSVHNNPQAGGSMRSTLFFNFGSMFVTDQAHKNEQWDEENLHRDEEYAS 414
D 302 VDPLVTLFGSVHNNPQAGGSMRSTLFFNFGSMFVTDQAHKNEQWDEENLHRDEEYAS 360
QY 415 DGAGDYEDNLHSPILSRQATGAEGKDI-VHGHGSGSALSMRRTLLGEGGDSSTDI 472
D 361 DHGADDIEDNLHSPILSRQATGAEGKDI-VHGHGSGSALSMRRTLLGEGGDSSTDI 415
QY 473 GGGWLAWKWESEGENGRKEGFKRVLHQRGVPGSRGSIIVSLPGGDDVPEGSEFVHAA 532
D 416 GGGWLAWKWESEGENGRKEGFKRVLHQRGVPGSRGSIIVSLPGGDDVPEGSEFVHAA 474
QY 533 AALYSQSALFSGKLAPRMSDAMVHPSEVAAKGRWKDLFEPGVRRALLVGVGIILOQ 592
D 475 AALYSQSALFSGKLAPRMSDAMVHPSEVAAKGRWKDLFEPGVRRALLVGVGIILOQ 534
QY 593 FAGINGVLYTTPQILBOAGVAVILSKFGLSSASASILISLTLMLPCIGFAMLLMDLS 652
D 535 FAGINGVLYTTPQILBOAGVAVILSKFGLSSASASILISLTLMLPCIGFAMLLMDLS 594
QY 653 GRFLLGTIPILIASLVLVNSLIDLTALHALLSTVSVIYCCFVWFGPIPNILCAE 712
D 595 GRFLLGTIPILIASLVLVNSLIDLTALHALLSTVSVIYCCFVWFGPIPNILCAE 654
QY 713 ABIEPTVRGICIAICAFWIGDIIIVTSLPMLNAIGLAGVSTIYAVVCLISFVFLKV 772
D 655 ABIEPTVRGICIAICAFWIGDIIIVTSLPMLNAIGLAGVSTIYAVVCLISFVFLKV 714
QY 773 KVPETKGMPLVITEFFAVGAKQA 795
D 715 KVPETKGMPLVITEFFAVGAKQA 737
RESULT 6
Q8LPQ8 ID Q8LPQ8 PRELIMINARY; PRT; 739 AA.
AC Q8LPQ8;
DT 01-OCT-2002 (Tremblrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE AT4935300/P23E12.140.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL: AY094465; AAM19835.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0005351; F:sugar porter activity; IEA.
DR GO: GO:0005215; F:transporter activity; IEA.
DR GO: GO:0008643; P:carbohydrate transport; IEA.
DR InterPro: IPR007114; MFS.
DR InterPro: IPR005828; Sub transporter.
DR InterPro: IPR003663; Sugar transport.
DR InterPro: IPR005829; Sug transporter.
DR Pfam: PF00083; Sugar tr; 1.
DR PRINTS: PR00171; SUGRTNSPOT.
DR PROSITE: PS00850; MFS; 1.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 739 AA; 79725 MW; 5D67C6D2DE7563FD CRC64;

Query Match 65.1%; Score 2654; DB 2; Length 739;
Best Local Similarity 70.0%; Pred. No. 5.4e-164;
Matches 521; Conservative 90; Mismatches 123; Indels 10; Gaps 7;

QY 56 MSGAVLVAIVASIGNLQGDWNTAAALVYIKKEFQLONEPTVEGLIVMSLIGATIVT 115
DB 1 MSGAVLVAIAAAGNLLQGDWNTAIAAGVYIKKEFNLSNPSEGLIVAMSLIGATLIT 60
QY 116 TFGSLSDSIRRRPMLLSILYFPGSLIMWSPNVYVLLARFVDFGGLAVTLVPLY 175
DB 61 TCSGGVADLGRPRMLLSILYFVGSILWLSNPVYVLLIGRLLDGFGVGLVTVLPYI 120
QY 176 ISETAPSEIRGLNLTLPQFGSGGMFLSYCMVFCWSLSPSPDMZIMLGVLAIPLRFFGL 235
DB 121 ISETAPPEIRGLNLTLPQFTSGGMFLSYCMVFCWSLSPSPDMZIMLGVLAIPLRFFGL 180
QY 236 TIFVLPSPRLVSKGRMAEAKVQLKRLGKDDVSGELSLLEGLVEGDTSEIEYIIGP 295
DB 181 TVFLPSPRLVSKGRMLEAKVQLRGLREDVSGEMALLVEGLIGIGETIEEYIIGP 240
QY 296 ATEAADDLVTDGKEQITLYCPBEGQGWIAAPSKPIMLSVLSLRHGS-WYNSQVPL 354
DB 241 ADEVTDHDIADVDDQIKLYGAEEGLSWAPVPVG---GSTMSVLSRRHGSTMSRRGSL 296
QY 355 MDPIVTLFGSVHNNPQAGGSMRSTLPFNFGSMFSTVDQIAKNPQWDEENLHRDDEYAS 414
DB 297 IDPLVTLFGSVHEKMPDPT-GSMRSLPFGSMFSGVGNQPRHEDWDEENLVGSGEDYPS 355
QY 415 DGAGDYEDNLHSLPLSQATGASKDQIVHHHGRGSLSMRQVLLGEGGDSVSTDIGG 474
DB 356 D-HGDDSDDLHSLPLSRQTSM-E-KMPHTAHGTLSFRGSGVQGAQEGAGSMGIGG 413
QY 475 GQWLAKWSEKEGNGRKEGKQVYLHOGVPGSRGSIYSLPGGSDVFGSGFVHAAA 534
DB 414 GQVAVKWKTEREDESQGEKQKRIYLYHOGVPGSRGSIYSLPGGSDTGE-ADFVQASA 472

QY 535 LVQSALFSKGLABPRMSDAAWHVPSVAAGSRWKOLFEPGVRRALLVGVGTOILOQFA 594
DB 473 LVSPALYSKDLLKHEHTIGPAMVHPS-ETTKSGIWHDLHPGVKRALWVGVLQIQQFS 531
QY 595 GINGVLYVTTPQLQAGVAVILSKFGLSSASASTILSSLTLLMLPCIGFAMLLMDLSGR 654
DB 532 GINGVLYVTTPQLQAGVAVILSKFGLSSASASTILSSLTLLMLPAIAVAMRLMDLSGR 591
QY 655 RFLLLGTIPILIASLVILVSNLIDLTGLAHALLSTVSVVYFCFVNGFGPIPNILCAE 714
DB 592 RTLLLTIPILIASLVILVSNLIDLTGLAHALLSTVSVVYFCFVNGFGPIPNILCAE 651
QY 715 IFFTRVRLCLTAICAFWIGDIIVTYSLPVMAIAGLAGVFSIYAVVCLISFVFEVLKV 774
DB 652 IFFTRVRLCLTAICAFWIGDIIVTYSLPVMAIAGLAGVFSIYAVVCLISFVFEVLKV 711
QY 775 PETKGMPLVITEFFFAVGAQAAA 798
DB 712 PETKGMPLVITEFFFAVGAQAAA 735

RESULT 7
O65497 PRELIMINARY; PRT; 729 AA.
ID O65497
AC O65497
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative sugar transporter protein (Monosaccharide sensing protein
DE 2).
GN Name=F23E12.140; Synonyms=AT4935300, mssp2;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,
RA Hohnselt J., Mewes H.W., Mayer K.F.X., Schueller C.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP EU Arabidopsis sequencing project;
RA Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Stamme C., Tjaden J., Trentmann O., Emmerlich V., Neuhaus E.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL: AL022604; CAAL8739.1; -.
DR EMBL: AL161587; CAB80247.1; -.
DR EMBL: AJ532570; CAD58692.1; -.
DR PIR: T06127; T06127.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0005351; F:sugar porter activity; IEA.
DR GO: GO:0005215; F:transporter activity; IEA.
DR GO: GO:0008643; P:carbohydrate transport; IEA.
DR InterPro: IPR007114; MFS.
DR InterPro: IPR005828; Sub transporter.
DR InterPro: IPR003663; Sugar transport.
DR InterPro: IPR005829; Sug transporter.

DR Pfam; PF00083; Sugar tr; 1.
 DR PRINTS; PRO0171; SUGTRNSPORT.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Sugar transporter; Transmembrane; transporter.
 SQ SEQUENCE 729 AA; 78524 MW; 81A35DDF8725411E CRC64;

Query Match 63.4%; Score 2584; DB 2; Length 729;
 Best Local Similarity 68.8%; Pred. No. 1.9e-159;
 Matches 512; Conservative 89; Mismatches 123; Indels 20; Gaps 8;

QY 56 MSGAVLVAIVASIGNILQGDNDNATIAAALVLYIKKEFOLQNEPTVEGLIVMSLIGATIVT 115
 DB 1 MSGAVLVAIAAAGNLLQGDNDNATIAAGVLYIKKEFNLSPNVEGLIVAMSLIGATLIT 60
 QY 116 TFGSGLSDSGRRPMLILSSILYFSGILMLSPNVYVLLARFVDFGFGIGLAVTLVPLV 175
 DB 61 TCSGGVADWLRPMLILSSILYFVGLVLMSPNVYVLLGLLDGFGVGLVTLVPIY 120
 QY 176 ISETAPEIRGLNLTLPQFGSGGMFLSYCMVFGMSLSPDPWRIMLGVLAIPSLPFFGL 235
 DB 121 ISETAPEIRGLNLTLPQFGSGGMFLSYCMVFGMSLSPDPWRIMLGVLAIPSLPFFGL 180
 QY 236 TIFVLPSPRLVSKGRMAAKVLQKRGKDDVSGELSLLRGLVGGDSTIEEYIIGP 295
 DB 181 TIFVLPSPRLVSKGRMAAKVLQKRGKDDVSGELSLLRGLVGGDSTIEEYIIGP 240
 QY 296 ATEAADPLVTDGKXEQITLYPPEGOSWIARPSKGPIMLGSVLASLRHGS-MVNGSVPL 354
 DB 241 ADEVTDHDIADVADKQIKLYAGEEGLSWARPVKG---GSTSVLSRHGSTMRSRQGS 296
 QY 355 MDPIVTLFGSVHNPQAGSMESTLFPNFGSMFSTVDQKAKNEQNDENLHRRDDREYAS 414
 DB 297 IDPLVTLFGSVHNPQAGSMESTLFPNFGSMFSTVDQKAKNEQNDENLHRRDDREYAS 355
 QY 415 DGAGGDDYEDNLHPLSLRQATGAEGKDIVHHGHRGSLSMRRQTLGEGGDGVSSTDIGG 474
 DB 356 D-HGDDSEDDLHPLSLRQATGAEGKDIVHHGHRGSLSMRRQTLGEGGDGVSSTDIGG 413
 QY 475 GWLAKWSEKGENRKEGKFRVYLHOGVPGSRGSIIVSLPGGDDVFEGSEFVHAAA 534
 DB 414 GWQVAMKWTREDESGOKE-----EGFPGSRGSIIVSLPGGDDGTGE-ADVFQASA 462
 QY 535 LVSGSALFSGLAEPMSDAAWHPSEVAAGSRWDLFEPGVRRALLVGVGIQIILQCPA 594
 DB 463 LVSGSALFSGLAEPMSDAAWHPSEVAAGSRWDLFEPGVRRALLVGVGIQIILQCPA 521
 QY 595 GINGVLYYTPQILEQAGVAVILSKFGLSASASILISITTLMLPFCIFAMLMDSGR 654
 DB 522 GINGVLYYTPQILEQAGVAVILSKFGLSASASILISITTLMLPFCIFAMLMDSGR 581
 QY 655 RELLIGTIPILIASLVLVNLDLGTALHALLSTVSVIVVFCFVWVGFGIPNLTCAE 714
 DB 582 RTULLTIPILIASLVLVNLDLGTALHALLSTVSVIVVFCFVWVGFGIPNLTCAE 641
 QY 715 IPFTRVGLCIAICAFPTWIGDIIVTYSPLVMLNAIGLAGVFSIYAVVGLISFVFVFLK 774
 DB 642 IPFTRVGLCIAICAFPTWIGDIIVTYSPLVMLNAIGLAGVFSIYAVVGLISFVFVFLK 701
 QY 775 PETKGMPELVITEFFRANGAKQAAA 798
 DB 702 PETKGMPELVITEFFRANGAKQAAA 725

RESULT 8

Q9SYQ3
 ID Q9SYQ3 PRELIMINARY; PRT; 734 AA.
 AC Q9SYQ3;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Sugar transporter.

GN Name=F0H16.18;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 NCBI_TaxID=3702;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.P.,
 RA Altieri H., Araujo R., Ruizar L., Rowley D., Buehler E., Dunn P.,
 RA Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,
 RA Luroso S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,
 Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -!- SIMILARITY: Belongs to the sugar transporter family.
 DR EMBL; AC007369; AAD30608.1; -;
 DR PIR; H86340; H86340.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005351; F:sugar porter activity; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0008643; P:carbohydrate transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub_transporter.
 DR InterPro; IPR003663; Sugar_transpt.
 DR InterPro; IPR005829; Sug_transporter.
 DR Pfam; PF00083; Sugar tr; 1.
 DR PRINTS; PRO0171; SUGTRNSPORT.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Sugar transporter; Transmembrane; transporter.
 SQ SEQUENCE 734 AA; 79425 MW; AB027864F850259C CRC64;

Query Match 55.9%; Score 2278; DB 2; Length 734;
 Best Local Similarity 61.2%; Pred. No. 1.6e-139;
 Matches 462; Conservative 112; Mismatches 145; Indels 36; Gaps 14;

QY 56 MSGAVLVAIVASIGNILQGDNDNATIAAALVLYIKKEFOLQNEPTVEGLIVMSLIGATIV 114
 DB 1 MKGATLVALAATIGNFLQGDNDNATIAAMVYIKNLNL---PTSVQGLVVMASLIGATVI 57
 QY 115 TTFSGPLSDSGRRPMLILSSILYFSGILMLSPNVYVLLARFVDFGFGIGLAVTLVPL 174
 DB 58 TTCGSPISDMGLRPMILSSVWYFVCGILMLSPNVYVLCFARLLNGFAGLAVTLVPV 117
 QY 175 YISIAPEIRGLNLTLPQFGSGGMFLSYCMVFGMSLSPDPWRIMLGVLAIPSLPFFG 234
 DB 118 YISIAPEIRGLNLTLPQFGSGGMFLSYCMVFTMSLSDSPSWRAMLGLVLSPLLYLF 177
 QY 235 LTIFVLPSPRLVSKGRMAAKVLQKRGKDDVSGELSLLRGLVGGDSTIEEYIIG 294
 DB 178 LTIFVLPSPRLVSKGRMAAKVLQKRGKDDVSGELSLLRGLVGGDSTIEEYIIG 237
 QY 295 PATEAADPLVTDGKXEQITLYPPEGOSWIARPSKGPIMLGSVLASLRHGSVMNQSV 353
 DB 238 LEDHEGDDTLTVDGKXEQITLYPPEGOSWIARPSKGPIMLGSVLASLRHGSVMNQSV 294
 QY 354 LMDPIVTLFGSVHNPQAGSMESTLFPNFGSMFSTVDQKAKNEQNDENLHRRDDREYAS 408
 DB 295 LKDPVLNLFGLSHKEMPEAGNTRSGIFPHFGSMFSTVDQKAKNEQNDENLHRRDDREYAS 354
 QY 409 DEEYASDAGGDYED---NLHSPLLSRQATGAEGKDIVHHGHRGSLSMRRQTLGEGGD 465
 DB 355 NDDYATDDGAGDDDDNDLRSPLMSRQT---SMD-KDMIPHTSGTSLSMRRHSTLMQ-GN 412
 QY 466 GVSTDTIGGGMQLANKWSEKGENRKEGKFRVYLHOGVPGSRGSIIVSLPGGDDVFEG 525
 DB 413 GESSMIGGGWHMGYRYENDE-----YKRYLKEDCAE-SRREGSIISIPGGPD--G 460
 QY 526 GSEFVHAAALVQSALFSGLAEPMSDAAWHPSEVAAGSRWDLFEPGVRRALLVGV 585
 DB 461 GGSVIAHASALVSRSLGPKS-----VHGSAMVPEKIAASGPLWSALLPFGVKRALVGV 515

586 GICILQFAGINGVLYTTPQILEQAGVAVILSKFGLSSASALISLSTLTLLMLPCIGFA 645
516 GICILQFAGINGVLYTTPQILERAGVDILSLGLSSISASFLISGLTLLMLPAIYVA 575
646 MLLMDLSGRRLFLGTPILIASLTVLVSNDIDITLAHALISTVSVIVVYFCFFVMGFG 705
576 MRLMDVSGRRSLLTWTPVLIVSLVAVISELTHISKVYNAALSTGCVLYVYFCFFVMGFG 635
706 PIPNLCARIEFPRVRGLCAICAFWIGDIIIVTSLPMLNAIGLAGVFSIYAVVCLI 765
636 PIPNLCSEIFPRVRGLCAICAMFWIGDIIIVTSLPVLSSIGLVGVFSIYAAVCVI 695
766 SFVFFVLKVPETKGMPLVITEFFAVGAKQAANKA 800
696 SWIFVYMKVPETKGMPLVITEFFAVGAKQAANKA 729

RESULT 9
Q9LNM67 PRELIMINARY; PRT; 734 AA.
AC Q9LNM67
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE F2D10.36.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Altafi H., Bei B., Chin C., Chioi J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Shinn P., Brooks S., Buehler E., Chao Q., Cheuk R., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chioi J., Choi E., Conn L.,
RA Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR EMBL; AC069251; AF080627.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transp.
DR InterPro; IPR005829; Sugar_transporter.
DR Pfam; PF00083; Sugar_tr; 1.
DR PRINTS; PR00171; SUGRTNSPORT.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 734 AA; 79484 MW; 0232FE8BF850259F CRC64;

Query Match 55.8%; Score 2275; DB 2; Length 734;
Best Local Similarity 61.28; Pred. No. 2.4e-139;
Matches 462; Conservative 111; Mismatches 146; Indels 36; Gaps 14;

56 MSGAVLVAIVASIGNLLQGDWNTAAAVLYIKKEFQLONEPT-VEGLIVSMISIGATIV 114
1 MKGATVALAATIGNFIQGDWNTAAGAVVYINKDLNL--PTSVQGLVWVMSLIGATVI 57
115 TTFSGPLSDSISGRPMILILSILYFSGILMNSPNVYVLLARFVDGFGGLAVTLVPL 174
58 TTCGPISDWIGRRPMILILSSVMYFVCGILMNSPNVYVLCFARLLNGFAGLAVTLVPV 117
175 YISIAIASEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPDWIRIMLGLVLAIFSLPFFG 234
118 YISETAPEIRGLNLTLPQFLGSGMFLSYCMVFTMSLSDSPSWRAMLGVLSIFSLLYLF 177
235 LTIFYPESPRMLVSKGRMABAKVLOKLRGKDVSGELSLLEGLVGGDTSTEEVIIG 294
178 LTFYLPESPRMLVSKGRMDEAKVLOQLCGREDVTDEALLVGEGLIGGKTMEDLLVT 237
295 PATEAADDLVTDGKE-QITLYGPEEQSWIARPSKGPIMLGSVLSLASRHGSMVNSVP 353
238 LEDHEGDDTLTVDEDDGMRLYGTENQSYLARVPPEQ--NSSLGLSRHSGSLANQSMI 294
354 LMDPIVTLFGSVHENMPOAGSMSTLFPNPGSMFVTDQ--HAKNQWD---BENLHRD 408
295 LKDPLVNLFGSLHEKMPAGNTRSGIFPHEGSMFTTADAPHGKPAHWEKDIESHYNKD 354
409 DEEVASDAGGDDYED--NLHSPILSRQATGAECKDIVHGHGHSALSMRROTLLGEGD 465
355 NDDVATDDGAGDDSDNDLSPMSRQTTSMD-KDMPHTSSTLSMRHSLTMQ-GN 412
466 GVSSTDIQGGWQLAWKSEKGEGRKGGPKRYLHQEGVPGRRGSIIVSLPGGDDVFE 525
413 GESSWGIGGWHMGVRYENDS-----YKRYLKEDEGAE-SRRGSIISIPGGPD--G 460
526 GSEFVHAALVQSALFSKGLAEPRMSDAAMVHSEVAAGSRWKLPEPGVRALLVGV 585
461 GGSVHASALVSRSVLGPKS-----VGSVAWPEKIAASGLPWSALLEPGVKALVGV 515
586 GICILQFAGINGVLYTTPQILEQAGVAVILSKFGLSSASALISLSTLTLLMLPCIGFA 645
516 GICILQFAGINGVLYTTPQILERAGVDILSLGLSSISASFLISGLTLLMLPAIYVA 575
646 MLLMDLSGRRLFLGTPILIASLTVLVSNDIDITLAHALISTVSVIVVYFCFFVMGFG 705
576 MRLMDVSGRRSLLTWTPVLIVSLVAVISELTHISKVYNAALSTGCVLYVYFCFFVMGFG 635
706 PIPNLCARIEFPRVRGLCAICAFWIGDIIIVTSLPMLNAIGLAGVFSIYAVVCLI 765
636 PIPNLCSEIFPRVRGLCAICAMFWIGDIIIVTSLPVLSSIGLVGVFSIYAAVCVI 695
766 SFVFFVLKVPETKGMPLVITEFFAVGAKQAANKA 800
696 SWIFVYMKVPETKGMPLVITEFFAVGAKQAANKA 729

RESULT 10
Q9SD00 PRELIMINARY; PRT; 729 AA.
AC Q9SD00
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE Sugar transporter-like protein (Monosaccharide sensing protein
3).
GN Name=F26O13.130; Synonyms=msp3;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Delsen M., Berger C., Cooke R., Grellert F., Laudie M., Mewes H.W.,
RA Lemcke K., Mayer K.P.X., Quetier F., Salanoubat M.;

RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;

RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA Scammie C., Tjaden J., Trentmann O., Emmerlich V., Neuhaus E.;

RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).

CC -1- SIMILARITY: Belongs to the sugar transporter family.

DR EMBL; AL133452; CAB63013.1; -

DR EMBL; AJ532571; CAD58693.1; -

DR PIR; T45780; T45780.

DR GO; GO:0016021; C: integral to membrane; IEA.

DR GO; GO:0005351; F: sugar porter activity; IEA.

DR GO; GO:0005215; F: transporter activity; IEA.

DR GO; GO:0008643; P: carbohydrate transport; IEA.

DR InterPro; IPR007114; MFS.

DR InterPro; IPR005828; Sub_transporter.

DR InterPro; IPR003663; Sugar_transp.

DR InterPro; IPR005829; Sug_transporter.

DR Pfam; PF00083; Sugar_tr; 1.

DR PRINTS; PR00171; SUGTRNSPORT.

DR PROSITE; PS00850; MFS; 1.

DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.

DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.

KW Sugar transport; Transmembrane; Transport.

SQ SEQUENCE 729 AA; 78828 MW; 241A9415DE43FCCC CRC64;

Query Match 55.1%; Score 2245; DB 2; Length 729;

Best Local Similarity 60.4%; Pred. No. 2.2e-137;

Matches 457; Conservative 110; Mismatches 144; Indels 46; Gaps 14;

QY 56 MSGAVLVAIVASTGNLQGDNDNATIAAAYIKKEFQIQNEPTVEGLIVMSLIGATIVT 115

DB 1 MRSVVLVALAAAGNMLQGDNDNATIAAGAVIYIKKEFHEKEPIEGLIVAMSLIGATIT 60

QY 116 TFGSLSDSTGRPMILASILYFEGSLIMLSPNVYLLARFVDGFGIGLAVTLVPLY 175

DB 61 TFGSPVSDKVRKSMILSGVLFSLVNFSPNVYLLFARLDGFGIGLAVTLVPLY 120

QY 176 ISEIAPSEIRGLNTLPQFGSGGMFLSCWYFGMSLSPDWRIMGLVLAIPSLFFPGL 235

DB 121 ISETAPSEIRGLNTLPQFGSGGMFLSCWYFGMSLSPDWRIMGLVLAIPSLFFPGL 180

QY 236 TIYLPESPRWLYSGRMAEAKVQLKRGKDDVSGELSLLEGVGGDTISIEYIIGP 295

DB 181 AAFLESPRWLYSGRMAEAKVQLKRGKDDVSGELSLLEGVGGDTISIEYIIGP 240

QY 296 ATEAADDVTDGDKGQITLYGPEEGSWIARPSKGPIMGLSVLSASRHGSMVNSVPLM 355

DB 241 DNEENEGNELPRKQIKLYGPDGSGWMAKPVKGQ----SSLALASRQGMPLPRGGSIM 296

QY 356 DPVITLFGSVHNP--QAGGSRNLTLPNFSGMSVTDQAHAKNQWDEENLHRDDEYA 413

DB 297 DPLVTLFGSIHENLPSNNMNSRSLFPNMGSLGMMGR--QESQWDPE---RNED-- 349

QY 414 SDGAGGDYDNLHSLPSSRQATGAECQIVHGHGSGALSMMRQTL--LQEGGDGVSSTDI 472

DB 350 ----SSDQDENLSPSSQT--EPDD--YHQRVTGTHNRQSSLPWANVGETATATSI 401

QY 473 GGGQWLAWKSWKEBNGRK--EGGFKRVYLHGE-----GVPGRRSGIVSL--PGGDV 523

DB 402 GGGQWLAWKYNKVGADGRVNGGLQRMVTHEETANNNTNTPFRRGSLLSFHPGEGDH 461

QY 524 FECSFEVHAALVSQALFSKGLABRMSDAAMVHPSEVAAGSRKWDLPFGVGERALLV 583

DB 462 DQVGVVQAALVSQAMVPGKGETML-----PKEV--KDGPGWRELKEPGVCRALMV 514

QY 584 GVGIQILOQFAGINGVLYTTPQILEQAGVAVILSKFGLSSASASLISLTLTLLMPCIG 643

DB 515 GVGLQILOQFAGINGVLYTTPQILEQAGVAVILSKFGLSSASASLISLTLTLLMPCIG 573

QY 644 FAMLMDLSGRFLLGLTIPILASLVLVVSNLDIGTLAHLSTVSVVVFCCFWMG 703

DB 574 -----LVSMRSLMLSTIPILLSLVTLVIGSLVNLGSGINALISTASVTVVLSCFWMG 626

QY 704 RGPINILCAEIPFTRVGRGLCAICAFTHWIGDIIVTVYSLPVMNLATGLAGVPSIYAVC 763

DB 627 FGAIPNLCSEIFPTSVRGICITICALTFCIDIVTVYTLVPMKLSIGIAGVFGIYAVC 686

QY 764 LISVFFVFLKVPETKGMPLVETIFFFAVGAQAQAAA 800

DB 687 AVANVFLKVPETKGMPLVETIFFFAVGAQAQAAA 723

RESULT 11

Q8GT51 PRELIMINARY; PRT; 753 AA.

AC Q8GT51; 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Sugar transporter.

GN Name=stp2;

OS Hordeum vulgare (Barley).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;

OC Triticeae; Hordeum.

OX NCBI_TaxID=4513;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Developing caryopsis;

RX MEDLINE=22424051; PubMed=12535352;

RA Weschke W., Panitz R., Gubatz S., Wang Q., Radchuk R., Weber H.,

RA Wobus U.;

RT "The role of invertases and hexose transporters in controlling sugar

RT ratios in maternal and filial tissues of barley caryopses during early

RT development.";

RL Plant J. 33:395-411(2003).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).

CC -1- SIMILARITY: Belongs to the sugar transporter family.

DR EMBL; AJ344446; CAD58959.1; -

DR GO; GO:0016021; C: integral to membrane; IEA.

DR GO; GO:0005351; F: sugar porter activity; IEA.

DR GO; GO:0005215; F: transporter activity; IEA.

DR GO; GO:0008643; P: carbohydrate transport; IEA.

DR InterPro; IPR007114; MFS.

DR InterPro; IPR005828; Sub_transporter.

DR InterPro; IPR003663; Sugar_transp.

DR InterPro; IPR005829; Sug_transporter.

DR Pfam; PF00083; Sugar_tr; 1.

DR PRINTS; PR00171; SUGTRNSPORT.

DR PROSITE; PS00850; MFS; 1.

DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.

KW Sugar transport; Transmembrane; Transport.

SQ SEQUENCE 753 AA; 81489 MW; A5332B341922A9B CRC64;

Query Match 53.8%; Score 2191.5; DB 2; Length 753;

Best Local Similarity 66.2%; Pred. No. 6.8e-134;

Matches 440; Conservative 76; Mismatches 132; Indels 17; Gaps 9;

QY 137 LYFSGILMLSNVYVLLARFVDGFGIGLAVTLVPLYISETAPSEIRGLNTLPQFS 196

DB 65 LVFLASNMVLAQNVYVLLARLDGFGIGLAVTLVPLYISETAPDTRGLNTLPQFS 124

QY 197 SGMFLSYCYMVGMSLSPDWRIMGLVLAIPSLFFGLTIFLPSPRWLYSGRMAEA 256

DB 125 SGMFLSYCYMVGMSLSPDWRIMGLVLAIPSLFFGLTIFLPSPRWLYSGRMAEA 184

QY 257 KTVLQKRGKDDVSGELSLLEGVGGDTISIEYIIGPATEADDLVTDGKEQITLYG 316

DB 185 KRVLQKRGKDDVSGELSLLEGVGGDTISIEYIIGPATEADDLVTDGKEQITLYG 242

QY 317 PERGQSWIARPSK--GPIMGLSVLSASRHGSMVNSVPLMDPIVTLFGSVHNPQAGG 374

Db 243 AEEGVSWIARVRGGQSALGALMSRHGSMVSGKSLVDPLVTLFGSVHEKPEVMG 302
 QY 375 SNESTLFPNFGSFSYTDQAKVQWDEENLHRDDDEYASDAGGDDYEDNLHSPILSRQA 434
 Db 303 SNESTLFPNFGSFSVAEQQAKADWDAES-HRDDDEYASDHGADDIEDNLNSPLSRQA 361
 QY 435 TGAEGKDIV-HHGRGSALSMMRQTLLEGGGVSSSTDIGGQWLAKWSEKENGGRK 492
 Db 362 TSVEGKEIAAPHGSIINGVSESSMQ-----GGDAVSSMGIGGQWLAKWTEREGADGRK 416
 QY 493 EGGFKVYLHQEPGSGRSISVSLPGGDDVFGSEFVHAALVSOALFSKGLAEPMS 552
 Db 417 EGGAQRIYLHEEBSGDRGSIUSMP-GGDIIPGGEYIQAAALVSPALYSKDLIEQQJA 475
 QY 553 DAAMVHPSEVAAGSRKWKLFEPGVRALLVGVGIQLQAFAGINVLVYTPQILEQAGV 612
 Db 476 GPAMVHPSEAVAKTRKAELEFEGVKHALFVGIGLQILQAFAGINVLVYTPQILEQAGV 535
 QY 613 AVLSKFGSSASASILISLTLMLPCIGFAMLMDSLSGRFLLGTLPTILIASLVIL 672
 Db 536 GILLSNIGLSSSSASILISALTLLMLGYISDRIGARAITAASFMYCSIPALFFVRRALL 595
 QY 673 VSNLIDLGTLAALLSTVSVIVVFCFVNGFGPIP--NILCAEIPPTVRVGLCIAICAF 730
 Db 596 VLVNVLDTGTVHAALSTSVIVVFCFVNM--GPIEGSTSSVKIITDRANAASLAKALL 653
 QY 731 TFWIGDIIVYTPVLMNAIGLAGVSIYAVVCLISFVFLVKVPTKGMPLVITEFEFA 790
 Db 654 TFWIGDIIVYTPVLMNAIGLAGVSIYAVVCMIAFVFFYMKVPTKGMPLVITEFEFS 713
 QY 791 VGAQK 795
 Db 714 VGAQK 718

RESULT 12
 Q96290
 ID Q96290 PRELIMINARY; PRT; 734 AA.
 AC Q96290;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Sugar transporter.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tjaden J., Neuhaus E.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -!- SIMILARITY: Belongs to the sugar transporter family.
 DR EMBL; Z50752; CAA90628.1; -;
 DR PIR; T51139; T51139.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005351; F:sugar porter activity; IEA.
 DR GO; GO:0003215; F:transporter activity; IEA.
 DR GO; GO:0008643; F:carbohydrate transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub transporter.
 DR InterPro; IPR003663; Sugar transp.
 DR InterPro; IPR005829; Sug transporter.
 DR Pfam; PF00083; Sugar tr_1.
 DR PRINTS; PR00171; SUGTRNSPRT.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR TRANSPORT 1; 1.
 DR PROSITE; PS00217; SUGAR TRANSPORT 2; 1.
 DR KX Sugar transport; Transmembrane; Transport.
 KW SEQUENCE 734 AA; 79462 MW; 1B009F52457AF6E1 CRC64;

Query Match 53.6%; Score 2183; DB 2; Length 734;
 Best Local Similarity 59.5%; Pred. No. 2.4e-133;
 Matches 449; Conservative 110; Mismatches 160; Indels 36; Gaps 14;
 QY 56 MSGAVLVAIVASIGNLLQGWDNATIAAAVLYIKKEFQLONEPT-VEGLIVSMISLIGATIV 114
 Db 1 MKGATLVALAATIGNFLQGWDNATIAAGAVYINKDLNL---PTSVQGLVWAMSLIGATIV 57
 QY 115 TTFSGPLSDSITGRPMILILSSILYFFSGLIMLWSPNYVLLARFVDFGFGIGLAVTLVPL 174
 Db 58 TTCGPIISDMIGRRPMLITLSSVYFVCLIMLWSPNYVLLCFARLLNGFGAGLAVTLVPL 117
 QY 175 YISETAPSEIEIGLNTLPQPSGSGMFLSYCMVFGMSLSPSPDWIRIMLVIAIPLSPFFFG 234
 Db 118 YISETAPSEIEIGLNTLPQFLSGSGMFLSYCMVFTXSLSDSPSNRMLGVLSISLLYLP 177
 QY 235 LTIYLPSPRLVSKGMAEAKVQLKLRGKDDVSGELSLLEGLVGGTSTIEEYIIG 294
 Db 178 LTVFLPSPRLVSKGMDKAEKRVLOQLCGREDVTDDEALLVEGLDIGGSKTMDLLVT 237
 QY 295 PATENADLVTDGKE-QITLYGPBEGOSWIARPSKGPIMLSGLSVLSASRHGSMVNOSVP 353
 Db 238 LEDHEGDDTLTETVEDDQIRLYGTHENQSYLARVPPEQ---NSSLGLRSRHGSLANQSMI 294
 QY 354 LMDPIVTLFGSVHENMPQAGSMRSTLFPNFGSMFVTDQ--HAKNEQWD---BENLHRD 408
 Db 295 LKDPLVNLFGSLHEKMPGAGNTRSGIFPHFGSMFSTTADAPHGKPAHWEKDIESHYNKD 354
 QY 409 DEEVASDAGGDYED---NLHSPILSRQATCAEGKDIVHGHGRGSALSMMRQTLLEGGD 465
 Db 355 NDDVATDDGAGDDDDSDNLSPLNSRQTTSDM-KDMIPHTSGTSLSMRRHSTLMQ-GN 412
 QY 466 GVSSTDIGGQWLAKWSEKEGNGRKGKRVLYHQEVGPGRRSGISVSLPGGDDVFE 525
 Db 413 GESSNGIGGGHMGVRYENDE-----YKRYLKEDGAE-SRSGSIISIPGGPD--G 460
 QY 526 GSEFVHAALVSOALFSKGLAEPMSDAAMVHPSEVAAGSRKWKLPPEGVRRALLVGV 585
 Db 461 GGSYIHASALVSRSVLGPKS-----VHGSAMVPEPKIAASGLPSALLEPGVKREALVGV 515
 QY 586 GTQLIQFAGINVLVYTPQILEQAGVAVILSKGLSSASASIISSLTLLMLPCIGFA 645
 Db 516 GTQLIQFSGINVLVYTPQILERAGVDILLSSGLSISASFLISGLTLLMLPAIWA 575
 QY 646 MLLMDLSGRFLLGTLPTILIASVLVSNLIDGLTAHALLSTVSVIVVFCFVNGFG 705
 Db 576 MELMDVSGRRSLLMTIPVLVSLVWLVISELHIHISKVNAALSTGCVLYFCFFVMGYG 635
 QY 706 PIPNLCIAEIPPTVRVGLCIAICAFTFWIGDIIVYTPVLMNAIGLAGVSIYAVVCL 765
 Db 636 PFQTSVLKSSQQAADRGGLCIAICAMVFWIGDIIVYTPVLSLSTLGVFVSIYAAVCVI 695
 QY 766 SFVFLKVPETKGMPLVITEFEFAVGAQAQAAKA 800
 Db 696 SWIFVYMKVPTKGMPLVITEFDYFAFGA-QAQA 729
 RESULT 13
 Q6K967
 ID Q6K967 PRELIMINARY; PRT; 652 AA.
 AC Q6K967;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Putative hexose transporter.
 GN Name=OJ1149_C12.19;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Sasaki T., Matsumoto T., Yamamoto K.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- SIMILARITY: Belongs to the sugar transporter family.
 DR EMBL; AP004082; BAD23011.1; -.
 DR InterPro; IPR000566; Lipocln_CytFABP.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub transporter.
 DR InterPro; IPR003663; Sugar_transpt.
 DR InterPro; IPR005829; Sug_transporter.
 DR Pfam; PF00083; Sugar_tr; 1.
 DR PRINTS; PR00171; SUGTRNSPT.
 DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 DR Sugar transport; Transmembrane; Transport.
 KW Sugar transport; Transmembrane; Transport.
 SQ SEQUENCE 552 AA; 68827 MW; EEE20446D2F9B1F6 CRC64;

 Query Match 48.2%; Score 1963; DB 2; Length 652;
 Best Local Similarity 55.1%; Pred. No. 4.2e-119;
 Matches 408; Conservative 95; Mismatches 134; Indels 104; Gaps 12;

 QY 56 MSGAVVAIVASIGNLQGDNDNATIAAALVIKKEF-QLNQPTVEGLIVMSLIGATIV 114
 DB 1 MRGAVAAAAAAGNNMLQGDNDNATIAALLDYMRRDLPALQHPALQGLVVAISLIGATIV 60

 QY 115 TTFSGLPSDSIGRRPMLIISILYFFSGILMSPNYYVLLARFVDGFGIGLAVTLVPL 174
 DB 61 TTFSGLPSDSRGRPRMLIASALLYSLAGLLMLSPNPVILLARLVDFGAIGLAVTLVPL 120

 QY 175 YSEIAPSETRGLNLTLPQSGSGGFLSYCMVFGMSLSPSPDWRIMGLVLAIPSLFFPG 234
 DB 121 YISETAPPTDTRGLNLTLPQSGTGMFLSYCMVFLITLAPINWRMLGLVLLPALLYL 180

 QY 235 LTIFYLPESPRVLVSGRMAEAKVKLQKRGDDVSGELSLLEGLVEGDDTSIEEYIG 294
 DB 181 LTIFYLPESPRVLVSGRMAEAKVKLQKRGDDVSGELSLLEGLVEGDDTSIEEYIG 240

 QY 295 PATEAADLVTDGDKBQITLYGPEEGQWIARPSKGPIMGLSVLSIASHSGMWNQSVPL 354
 DB 241 PTDEA-----KVTLYG---GMSSGLAPGS---MFGSAVTLASQGSMLDH---L 280

 QY 355 MDPIVTLFGSVHNPQAGSGMESTLFPNFGMFSVTDQAHKNEQWDEENLHRDDEEYAS 414
 DB 281 KDPVALLDSLDHNPAGG---TTDVPNLGSMIGVHDRPID--WDEEN----- 325

 QY 415 DGAGGYEDNLHPLLSRQATGAEGKDIVHGHRSALSMRRQTLLEGCGDGVSSDTIGG 474
 DB 326 ---SGDDGDIAPLLTM-----EGEAATSTVGIGG 353

 QY 475 GWOLANKWSEKEGKGGPKRVYLHOEGVPGSRGIVSLPGCGDVFESGFVHAAA 534
 DB 354 GWOLANKWTEGVAADGTROSTVKRMVYLHBEQAE-----VHAAA 392

 QY 535 LVQSALFSGKGLAEPKMSDAAMVHPSEVAAKGRKDKLPEP-GVRRALLVGVGIQLQOF 593
 DB 393 LVQSALCTKKEAEAEVEGG-----NREVLPGGVRRHALVCGVAIQILQOF 438

 QY 594 AGINGVLYTPQILEAGVAVILSKFGLSSASASILISLTLMLPCIGFANMLMDLSG 653
 DB 439 SGISGVLLTPQILEAGVGVLLSRGLRDDDSASILISGVITLMLPSIGVAMRLMDVSG 498

 QY 654 RPELLGTPILIASIVILVSNLIDLTALHALLSTSVIVYVFCFVNGFGPIENILCA 713
 DB 499 RRSLLWTLPIPLVASLAVLVAASVAPAAAHAAVCTGSVVYLCFFVNGFGPIENILCA 558

 QY 714 EIPPTVRGLCIAICFTWIGDIIVTYSLPVNLNAIGLAVFSYVAVCLISFVFLK 773
 DB 559 EIPPTVRGLCIAICSLAFWADIATVYTLVPLASGLAGLFAIYAAVCCVALVFAVR 618

 QY 774 VPETKMPLEVIETFFAVGAK 794
 DB 619 VPETKGLPLEVIIDFFNVGAK 639

RESULT 14
 Q8H887 PRELIMINARY; PRT; 714 AA.
 ID Q8H887;
 AC Q8H887;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE Putative sugar transporter protein.
 GN Name=OUI528D07.4;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=33947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wing R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Sasaki C.,
 RA Currie J., Collura K.;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- SIMILARITY: Belongs to the sugar transporter family.
 DR EMBL; AC099739; AAN17390.1; -.
 DR Gramene; Q8H887; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005351; F:sugar porter activity; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0008643; P:carbohydrate transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub_transporter.
 DR InterPro; IPR003663; Sugar_transpt.
 DR InterPro; IPR005829; Sug_transporter.
 DR Pfam; PF00083; Sugar_tr; 1.
 DR PRINTS; PR00171; SUGTRNSPT.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Sugar transport; Transmembrane; Transport.
 SQ SEQUENCE 714 AA; 77882 MW; E6DA98A692106885 CRC64;

 Query Match 43.3%; Score 1763; DB 2; Length 714;
 Best Local Similarity 47.5%; Pred. No. 4.8e-106;
 Matches 377; Conservative 125; Mismatches 193; Indels 98; Gaps 14;

 QY 4 GSWLAVPTPTDLDRLRLPSVLLALPGLPPASCSQEPVTSDDILEDKMGAVLVA 63
 DB 9 GAW-----KNTRQYSPSPISLGSVFSSESLSLSTTTAD---GSTANAVLAA 54

 QY 64 IVASIGNLLQGDNDNATIAAALVIKKEFQLNQPTVEGLIVMSLIGATIVTFSPGLSD 123
 DB 55 IVASIGNLLQGDNDNATIAAALVIKKEFNQNDPMVEGLILANSLIGATIITALSMTN 114

 QY 124 SIGRRPMLIISILYFFSGILMSPNYYVLLARFVDGFGIGLAVTLVPLYSIETAPSE 183
 DB 115 SIGRRPMLIISILYFFSGILMSPNYYVLLARLYGFGSLVTVYAPLYSETAPT 174

 QY 184 IRLGLNLTLPQSGGGMFLSYCMVFGMSLSPSPDWRIMGLVLAIPSLFFGLTIFVLPES 243
 DB 175 MRGLNLTLPQSGGGMFLSYCMVFGMSLSPSPDWRIMGLVLAIPSLFFGLTIFVLPES 234

 QY 244 PRLVSGRMAEAKVKLQKRGDDVSGELSLLEGLVEGDDTSIEEYIGPATEAADL 303
 DB 235 PRLVSGRMAEAKVKLQKRGDDVSGELSLLEGLVEGDDTSIEEYIGPATEAADL 293

 QY 304 VTDCRKEQITLYGPEEGQWIARPSKGPIMGLSVLSIASHSGMWNQSVPLMDPIVTEG 363
 DB 294 QSIENPTTKLYGHEEGVTWFAFPFGKNV-----VESDHSPIFN---LLDPIVTEG 343

 QY 364 SVIENPFAQGSRRSTLFPNFGMFSVTDQAHKNEQWDEENLHRDDEEYASDAGGDDY 423
 DB 344 SIHGN-----LNTPEFTSSGNMSNDIEQ-PKTDLESQEDLDT-----DYED 384

```

QY 424 NLHSPILLSROATGAEKDIIVHGHGSGALSMMRQTLLEGGDVSTDIIGGQWLAKWS 483
Db 385 DLGHELLF-----HQSVM-----EGIDDACVNGGWHIAKVF 417
QY 484 EKEGNGRKEGFKRVYLHOGVPGVRRGSIIVSLPGGVDFEGSEFVHAAALVSQALFS 543
Db 418 QRENEFGQTDQDFQIFLQ-----GDIQAGRVSHATALVSTPS-FH 458
QY 544 KGLAEPRMSDAAMVHPSE--VAAGSRWKDLFPFGVRRALLVGVG:QIILQOFAGINGVLY 601
Db 459 HSIG-----PAMVHPKFNLSGQSWSDLLQPGVKQGLIVGVT:QIILQLAGISGLY 512
QY 602 YTPQILEQAGAVILSKFGLSSASALISLTTLLMPCIGFAMLLMDLSGRRELLIGT 661
Db 513 YTPQILEQAGAGILLKWFNVSSSSSILTSALTFTFMLSIGIAMKMDRYGRSLLIYT 572
QY 662 IPILIASLVILVNSNIDLGTLAHLSTVSVIYFCFVNGGPIPNILCAEIPTRVR 721
Db 573 IPMLIVSLIILIVNMWNLAEIAGAILSPGVIYVCCFVNGGPIPNVLCSELFPFSCR 632
QY 722 GLCIAICAFWTGDIIVTVSLPMLNAIGLAGVFSIYAVVCLISFVFLKVPETKGM 781
Db 633 NRCMSICTLFTFWISVIIVTAFVPMSSIGLIGVCGIYAVVCIVSFIIVLIRKVPETKGM 692
QY 782 LEVITEFFAVGAK 794
Db 693 LAVIANSLAVGAR 705

RESULT 15
Q9SNK7 PRELIMINARY; PRT; 723 AA.
AC Q9SNK7
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Similar to sugar transporter protein.
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL; AP000615; BAA85398.1; -.
DR Gramene; Q9SNK7; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; Sugar tr; 1.
DR PRINTS; PR00171; SUGTRNSPOT.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 723 AA; 79061 MW; EC1E924FC4206BD8 CRC64;

```

Query Match 43.3%; Score 1763; DB 2; Length 723;
 Best Local Similarity 47.5%; Pred. No. 4.9e-106;
 Matches 377; Conservative 125; Mismatches 193; Indels 98; Gaps 14;

```

QY 4 GSWLAVTPTPTDRLRLLPSVVLALPGCLPPACSCSQEPVTSDILEDRMVGAVLVA 63
Db 18 GAW-----KNTRQYSPSPISLGSVFSSESSLELSTTTAD-----GSTANAVLAA 63

```

Search completed: October 13, 2004, 11:36:05
 Job time : 95.8744 secs

```

QY 64 IVASIGNLLOQWDNATIAAAVLYIKKFPOLQNEPTVEGLIVSMISLIGATIVTTSGPLSD 123
Db 64 IVASIGNLLOQWDNAAAGATIMYIKNEFNLDQPMMEGLILAMSLIGATITALSGMTIN 123
QY 124 SIGRRPMLILSILYFFSGGLIMLSPNVYVLLARFVDGFGIGLAVTIVPLVYISBIAPE 183
Db 124 SIGKRPILLSVAALYSISALIMFOASNEYMLLARLIVFGSGLVTVVAPLYISETAPT 183
QY 184 IRGLINTLPQPSGGGMFLSCVNFVGMSLSPSPDWRMLGVLAIPSLFFGLTIFLYPE 243
Db 184 MREGLINTLPQFNGSLGMLLSYIMVFLMSLTLPNWRMLGSLSPSFVFLLCIFLYPE 243
QY 244 PRWLVSCKRMAEAKVLKLRGKDDVSGELSLLEGLEVEGGDTIEEYIIIPATEAADL 303
Db 244 PVFLVSKKIEEAKVMKRLRGTVNEVSEIAFLIQLTVDQDNVIEDWIGHNDEFDD 302
QY 304 VTDGKKEITLYGPEGQSWIARPSKIPMLGSLVLSASRHGSMVNGSVPLMDPITVLF 363
Db 303 QSIISNTETTKLYGHEEGVTWEPFPGKNV-----VESDHPSPN-----LLDPITVLF 352
QY 364 SVHENWPAQSGMSRSTLPFNFGSMFVTDQHAKEQWDEENLHRDDEYASDGGADYED 423
Db 353 SIHGINI-----LNPTEFTSSGNMSNDIEQ-PKTDLESQEDLT-----DYED 393
QY 424 NLHSPILLSROATGAEKDIIVHGHGSGALSMMRQTLLEGGDVSTDIIGGQWLAKWS 483
Db 394 DLGHELLF-----HQSVM-----EGIDDACVNGGWHIAKVF 426
QY 484 EKEGNGRKEGFKRVYLHOGVPGVRRGSIIVSLPGGVDFEGSEFVHAAALVSQALFS 543
Db 427 QRENEFGQTDQDFQIFLQ-----GDIQAGRVSHATALVSTPS-FH 467
QY 544 KGLAEPRMSDAAMVHPSE--VAAGSRWKDLFPFGVRRALLVGVG:QIILQOFAGINGVLY 601
Db 458 HSIG-----PAMVHPKFNLSGQSWSDLLQPGVKQGLIVGVT:QIILQLAGISGLY 521
QY 602 YTPQILEQAGAVILSKFGLSSASALISLTTLLMPCIGFAMLLMDLSGRRELLIGT 661
Db 522 YTPQILEQAGAGILLKWFNVSSSSSILTSALTFTFMLSIGIAMKMDRYGRSLLIYT 581
QY 662 IPILIASLVILVNSNIDLGTLAHLSTVSVIYFCFVNGGPIPNILCAEIPTRVR 721
Db 582 IPMLIVSLIILIVNMWNLAEIAGAILSPGVIYVCCFVNGGPIPNVLCSELFPFSCR 641
QY 722 GLCIAICAFWTGDIIVTVSLPMLNAIGLAGVFSIYAVVCLISFVFLKVPETKGM 781
Db 642 NRCMSICTLFTFWISVIIVTAFVPMSSIGLIGVCGIYAVVCIVSFIIVLIRKVPETKGM 701
QY 782 LEVITEFFAVGAK 794
Db 702 LAVIANSLAVGAR 714

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 11:33:18 ; Search time 55.5861 Seconds
(without alignments)
3568.829 Million cell updates/sec

Title: US-10-051-909-36
Perfect score: 2779
Sequence: 1 PSSSSFRPAGKKKKKNOG.....TIFVLSIQRLQWLPECLUS 553

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A_Geneseq_23Sep04:*
- 1: Geneseq1980s:*
 - 2: Geneseq1990s:*
 - 3: Geneseq2000s:*
 - 4: Geneseq2001s:*
 - 5: Geneseq2002s:*
 - 6: Geneseq2003as:*
 - 7: Geneseq2003bs:*
 - 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2777	99.9	553	8	ADG47940 Corn Beta
2	1801.5	64.8	501	8	ADG47942 Beta-vulg
3	951.5	34.2	437	3	AAG13750 Arabidops
4	951.5	34.2	447	3	AAG13749 Arabidops
5	951.5	34.2	463	3	AAG13748 Arabidops
6	943.5	34.0	437	3	AAG20978 Arabidops
7	943.5	34.0	447	3	AAG20977 Arabidops
8	943.5	34.0	463	3	AAG20976 Arabidops
9	933	33.6	438	3	AAG38871 Arabidops
10	933	33.6	448	3	AAG38870 Arabidops
11	933	33.6	464	3	AAG38869 Arabidops
12	891	32.1	481	8	ADM48307 Polypepti
13	874.5	31.5	468	3	AAG3282 Arabidops
14	874.5	31.5	470	3	AAG3281 Arabidops
15	868.5	31.3	474	3	AAG50066 Arabidops
16	853.5	30.7	472	3	AAG50064 Arabidops
17	853.5	30.7	474	3	AAG50063 Arabidops
18	829.5	29.8	415	3	AAG50067 Arabidops
19	824	29.7	477	3	AAG30433 Arabidops
20	824	29.7	483	3	AAG30432 Arabidops
21	819	29.5	420	3	AAG30434 Arabidops
22	807.5	29.1	415	3	AAG50065 Arabidops
23	805.5	29.0	471	3	AAG48646 Arabidops
24	805.5	29.0	558	3	AAG48645 Arabidops
25	805.5	29.0	564	3	AAG48644 Arabidops

26	801.5	28.8	398	3	AAG39283 Arabidops
27	790.5	28.4	454	3	AAG37675 Arabidops
28	790.5	28.4	459	3	AAG37674 Arabidops
29	786.5	28.3	454	3	AAG06925 Arabidops
30	786.5	28.3	460	3	AAG06924 Arabidops
31	774.5	27.9	409	3	AAG37676 Arabidops
32	770.5	27.7	409	3	AAG06926 Arabidops
33	758.5	27.3	382	3	AAG10665 Arabidops
34	741.5	26.7	375	3	AAG10666 Arabidops
35	738.5	26.6	384	3	AAG50068 Arabidops
36	717.5	25.8	369	3	AAG10667 Arabidops
37	688	24.8	489	4	ABM63511 Drosophil
38	682.5	24.6	370	3	AAG03696 Arabidops
39	681.5	24.5	368	3	AAG03697 Arabidops
40	671	24.1	477	4	AAE66934 Murine GL
41	668.5	24.1	478	4	AAE66933 Rat GLUTX
42	661.5	23.8	478	4	AAE66939 GLUTX1 co
43	658.5	23.7	477	4	AAE04888 Human tra
44	658.5	23.7	477	6	ABP58364 Human sol
45	652.5	23.5	477	4	AAE66932 Human GLU

ALIGNMENTS

RESULT 1
ADG47940
ID ADG47940 standard; protein; 553 AA.
XX
AC ADG47940;
XX
DT 11-MAR-2004 (first entry)
XX
DE Corn Beta-vulgaris-like sugar transport protein #4.
XX
KW Arabidopsis-like sugar transport protein;
KW Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;
KW corn; plant.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT Misc-difference 528
FT /note= "Encoded by AGN"
XX
PN US2002199217-A1.
XX
PD 26-DEC-2002.
XX
PF 17-JAN-2002; 2002US-00051909.
XX
PR 24-APR-1998; 98US-0083044P.
PR 14-APR-1999; 99US-00291922.
XX
PA (HELE/) HELENTJARIS T G.
XX
PI Helentjaris TG;
XX
DR WPI; 2004-040967/04.
DR N-PSDB; ADG47939.
XX
FT New isolated polynucleotide encoding a polypeptide having sugar transport
FT protein activity, for producing a transformed plant and for use as probes
FT in physical mapping.
XX
PS Claim 18; SEQ ID NO 36; 71pp; English.
XX
CC The invention relates to Arabidopsis-like or Beta-vulgaris-like sugar
CC transport proteins and their corresponding nucleic acid sequences. The
CC sequences of the invention are useful to transform a cell. These are also
CC useful to produce a transgenic plant. Probes derived from sequences
CC encoding sugar transport protein may be used for physical mapping. The
CC present sequence is corn Beta-vulgaris-like sugar transport protein.

KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX Arabidopsis thaliana.
 OS EP1033405-A2.
 PN
 XX
 XX
 PD 06-SEP-2000.
 XX
 XX 25-FEB-2000; 2000EP-00301439.
 XX
 PR 25-FEB-1999; 99US-0121825P.
 PR 05-MAR-1999; 99US-0123180P.
 PR 09-MAR-1999; 99US-0123548P.
 PR 23-MAR-1999; 99US-0135788P.
 PR 25-MAR-1999; 99US-0138268P.
 PR 29-MAR-1999; 99US-0138788P.
 PR 01-APR-1999; 99US-0127463P.
 PR 06-APR-1999; 99US-0128234P.
 PR 16-APR-1999; 99US-0128714P.
 PR 19-APR-1999; 99US-0129845P.
 PR 21-APR-1999; 99US-0130077P.
 PR 23-APR-1999; 99US-0130449P.
 PR 23-APR-1999; 99US-0130504P.
 PR 23-APR-1999; 99US-0130891P.
 PR 28-APR-1999; 99US-0131449P.
 PR 30-APR-1999; 99US-0132048P.
 PR 30-APR-1999; 99US-0132407P.
 PR 04-MAY-1999; 99US-0132484P.
 PR 05-MAY-1999; 99US-0132485P.
 PR 06-MAY-1999; 99US-0132486P.
 PR 06-MAY-1999; 99US-0132487P.
 PR 07-MAY-1999; 99US-0132863P.
 PR 11-MAY-1999; 99US-0134256P.
 PR 14-MAY-1999; 99US-0134218P.
 PR 14-MAY-1999; 99US-0134219P.
 PR 14-MAY-1999; 99US-0134221P.
 PR 18-MAY-1999; 99US-0134370P.
 PR 18-MAY-1999; 99US-0134768P.
 PR 19-MAY-1999; 99US-0134941P.
 PR 20-MAY-1999; 99US-0135124P.
 PR 21-MAY-1999; 99US-0135363P.
 PR 24-MAY-1999; 99US-0135629P.
 PR 25-MAY-1999; 99US-0136021P.
 PR 27-MAY-1999; 99US-0136332P.
 PR 28-MAY-1999; 99US-0136782P.
 PR 01-JUN-1999; 99US-0137222P.
 PR 03-JUN-1999; 99US-0137528P.
 PR 04-JUN-1999; 99US-0137502P.
 PR 07-JUN-1999; 99US-0137724P.
 PR 08-JUN-1999; 99US-0138094P.
 PR 10-JUN-1999; 99US-0138540P.
 PR 10-JUN-1999; 99US-0138847P.
 PR 14-JUN-1999; 99US-0139119P.
 PR 16-JUN-1999; 99US-0139452P.
 PR 16-JUN-1999; 99US-0139453P.
 PR 17-JUN-1999; 99US-0139492P.
 PR 18-JUN-1999; 99US-0139454P.
 PR 18-JUN-1999; 99US-0139455P.
 PR 18-JUN-1999; 99US-0139456P.
 PR 18-JUN-1999; 99US-0139457P.
 PR 18-JUN-1999; 99US-0139458P.
 PR 18-JUN-1999; 99US-0139459P.
 PR 18-JUN-1999; 99US-0139460P.
 PR 18-JUN-1999; 99US-0139461P.
 PR 18-JUN-1999; 99US-0139462P.
 PR 18-JUN-1999; 99US-0139463P.
 PR 18-JUN-1999; 99US-0139750P.
 PR 18-JUN-1999; 99US-0139763P.
 PR 21-JUN-1999; 99US-0139887P.
 PR 22-JUN-1999; 99US-0139899P.
 PR 23-JUN-1999; 99US-0140363P.
 PR 23-JUN-1999; 99US-0140354P.
 PR 24-JUN-1999; 99US-0140695P.
 PR 28-JUN-1999; 99US-0140823P.
 PR 29-JUN-1999; 99US-0140991P.
 PR 30-JUN-1999; 99US-0141287P.
 PR 01-JUL-1999; 99US-0141842P.
 PR 01-JUL-1999; 99US-0142154P.
 PR 02-JUL-1999; 99US-0142055P.
 PR 06-JUL-1999; 99US-0142390P.
 PR 08-JUL-1999; 99US-0142803P.
 PR 09-JUL-1999; 99US-0142920P.
 PR 12-JUL-1999; 99US-0142977P.
 PR 13-JUL-1999; 99US-0143542P.
 PR 15-JUL-1999; 99US-0143624P.
 PR 16-JUL-1999; 99US-0144085P.
 PR 16-JUL-1999; 99US-0144086P.
 PR 19-JUL-1999; 99US-0144325P.
 PR 19-JUL-1999; 99US-0144331P.
 PR 19-JUL-1999; 99US-0144332P.
 PR 19-JUL-1999; 99US-0144333P.
 PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 21-JUL-1999; 99US-0144884P.
 PR 21-JUL-1999; 99US-0144814P.
 PR 21-JUL-1999; 99US-0145086P.
 PR 22-JUL-1999; 99US-0145085P.
 PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145152P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 26-JUL-1999; 99US-0145224P.
 PR 27-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145918P.
 PR 27-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145951P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 02-AUG-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 10-AUG-1999; 99US-0147935P.
 PR 11-AUG-1999; 99US-0148171P.
 PR 12-AUG-1999; 99US-0148319P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149929P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 25-AUG-1999; 99US-0149930P.
 PR 26-AUG-1999; 99US-0150566P.
 PR 27-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 30-AUG-1999; 99US-0151080P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.

PR	07-SEP-1999;	99US-0152363P.	QY	460	MMSEILPVSIKSLGSIATLANWLTSPAITMTNMLTWSVGGTFLSYMYSAFTIVFW	519
PR	10-SEP-1999;	99US-0153070P.	Db	358	VMSEIFPINIKGVAGGMATLVNWFCAWAVSYTFNFLMSWSSYGIFLIYAANALAIIVFI	417
PR	13-SEP-1999;	99US-0153758P.				
PR	15-SEP-1999;	99US-0154018P.				
PR	16-SEP-1999;	99US-0154039P.				
PR	20-SEP-1999;	99US-0154779P.	QY	520	LWVPETKG	527
PR	22-SEP-1999;	99US-0155139P.	Db	418	AIVPETKG	425
PR	23-SEP-1999;	99US-0155486P.				
PR	24-SEP-1999;	99US-0155659P.				
PR	28-SEP-1999;	99US-0156458P.				
PR	29-SEP-1999;	99US-0156596P.				
PR	04-OCT-1999;	99US-0157117P.				
PR	05-OCT-1999;	99US-0157753P.				
PR	06-OCT-1999;	99US-0157856P.				
PR	07-OCT-1999;	99US-0158029P.				
PR	08-OCT-1999;	99US-0158232P.				
PR	12-OCT-1999;	99US-0158369P.				
PR	13-OCT-1999;	99US-0159293P.				
PR	13-OCT-1999;	99US-0159294P.				
PR	13-OCT-1999;	99US-0159295P.				
PR	14-OCT-1999;	99US-0159328P.				
PR	14-OCT-1999;	99US-0159330P.				
PR	14-OCT-1999;	99US-0159331P.				
PR	14-OCT-1999;	99US-0159637P.				
PR	14-OCT-1999;	99US-0159638P.				
PR	18-OCT-1999;	99US-0159584P.				
PR	21-OCT-1999;	99US-0160741P.				
PR	21-OCT-1999;	99US-0160767P.				
PR	21-OCT-1999;	99US-0160768P.				
PR	21-OCT-1999;	99US-0160770P.				
PR	21-OCT-1999;	99US-0160814P.				
PR	21-OCT-1999;	99US-0160815P.				
PR	22-OCT-1999;	99US-0160980P.				
PR	22-OCT-1999;	99US-0160981P.				
PR	22-OCT-1999;	99US-0160989P.				
PR	23-OCT-1999;	99US-0161404P.				
PR	23-OCT-1999;	99US-0161405P.				
PR	23-OCT-1999;	99US-0161406P.				
PR	26-OCT-1999;	99US-0161359P.				
PR	26-OCT-1999;	99US-0161360P.				
PR	26-OCT-1999;	99US-0161361P.				
PR	28-OCT-1999;	99US-0161320P.				
PR	28-OCT-1999;	99US-0161992P.				
PR	28-OCT-1999;	99US-0161993P.				
PR	29-OCT-1999;	99US-0162142P.				
Query Match 34.2%; Score 951.5; DB 3; Length 437;						
Best Local Similarity 42.3%; Pred. No. 1.1e-79; Mismatches 140; Indels 5; Gaps 2;						
Matches 181; Conservative 102; Mismatches 140; Indels 5; Gaps 2;						
QY	100	FLCTLVALGPIQFGTSSFTQDAMVRDLNLISSEFSAFGSLNSVGGMVGAIASQGM	159			
Db	3	YLSFTFVAVCGPFAFGSCAGYSPAQAIRNDLSLITAEPSLFGSLTTCGAMIGALTSGP	62			
QY	160	AEVIGKSLMIAAIPIGMLAIKSPAKDASLYNCRLLGEGVGLIISYTPVYIAEISP	219			
Db	63	ADLVGRKGAHVSSAFCVGVGMLAIIPAKGVVLDLGRATGGMGAFSYVVPFIAEIAEP	122			
QY	220	QNRGALGVNQLSVTFGLFAYLLGMFIPIWRLLAIVGALPCTMLIPGLFFIPESPRMLA	279			
Db	123	KTFRGALTTNLQILICTGVSVSFIIGTLVTRVLAIGIIPCAASFGLGFFIPESPRMLA	182			
QY	280	KMNLTEDETSQVLRGPFTDITTEVNDIKRAVASSKRTTISFOELNKKYRTPLLGI	339			
Db	183	KVGRTEFEAALURKGRKADISEAAEQDIETLERLPKAKMLDLFQRYIRSVLIAF	242			
QY	340	GLLVQLNLSINGVLFIYASSIFKAAGVNTSDIATCSLGAIOVLATGVTTLLDRAGRIL	399			
Db	243	GLMVFQFGGINGICFYTSISFEQAGFP-TRGLMIYAVLQVITALNAPIVDRAGKPL	301			
QY	400	LIISGTMTCLLAVSVFFLKDNTISQDSNSYIITMISLGVISFVITPFGMGAPWL	459			
Db	302	LLVSATGLVIGLIIAAVSPYLKVH----DMAHEAVPVLAVGVIMVYIGSFAGMGAMPW	357			

QY	460	MMSEILPVSIKSLGSIATLANWLTSPAITMTNMLTWSVGGTFLSYMYSAFTIVFW	519
Db	358	VMSEIFPINIKGVAGGMATLVNWFCAWAVSYTFNFLMSWSSYGIFLIYAANALAIIVFI	417
QY	520	LWVPETKG	527
Db	418	AIVPETKG	425
RESULT 4			
AAG13749			
ID	AAG13749 standard; protein; 447 AA.		
XX	AAG13749;		
XX	17-OCT-2000 (first entry)		
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 13358.		
DE	Protein identification; signal transduction pathway; metabolic pathway;		
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		
KW	termination sequence.		
XX	Arabidopsis thaliana.		
XX	EP1033405-A2.		
XX	06-SEP-2000.		
XX	25-FEB-2000; 2000EP-00301439.		
PR	25-FEB-1999;	99US-0121825P.	
PR	05-MAR-1999;	99US-0123180P.	
PR	09-MAR-1999;	99US-0123548P.	
PR	23-MAR-1999;	99US-0125788P.	
PR	25-MAR-1999;	99US-0126264P.	
PR	29-MAR-1999;	99US-0126785P.	
PR	01-APR-1999;	99US-0127462P.	
PR	06-APR-1999;	99US-0128234P.	
PR	08-APR-1999;	99US-0128714P.	
PR	16-APR-1999;	99US-0129845P.	
PR	19-APR-1999;	99US-0130077P.	
PR	21-APR-1999;	99US-0130449P.	
PR	23-APR-1999;	99US-0130510P.	
PR	23-APR-1999;	99US-0130891P.	
PR	28-APR-1999;	99US-0131449P.	
PR	30-APR-1999;	99US-0132048P.	
PR	30-APR-1999;	99US-0132407P.	
PR	04-MAY-1999;	99US-0132484P.	
PR	05-MAY-1999;	99US-0132485P.	
PR	06-MAY-1999;	99US-0132486P.	
PR	06-MAY-1999;	99US-0132487P.	
PR	07-MAY-1999;	99US-0132863P.	
PR	11-MAY-1999;	99US-0134256P.	
PR	14-MAY-1999;	99US-0134218P.	
PR	14-MAY-1999;	99US-0134219P.	
PR	14-MAY-1999;	99US-0134221P.	
PR	18-MAY-1999;	99US-0134370P.	
PR	18-MAY-1999;	99US-0134768P.	
PR	19-MAY-1999;	99US-0134941P.	
PR	20-MAY-1999;	99US-0135124P.	
PR	21-MAY-1999;	99US-0135353P.	
PR	24-MAY-1999;	99US-0135629P.	
PR	25-MAY-1999;	99US-0136021P.	
PR	27-MAY-1999;	99US-0136392P.	
PR	28-MAY-1999;	99US-0136782P.	
PR	01-JUN-1999;	99US-0137222P.	
PR	03-JUN-1999;	99US-0137528P.	
PR	07-JUN-1999;	99US-0137502P.	
PR	08-JUN-1999;	99US-0138094P.	
PR	10-JUN-1999;	99US-0138540P.	

10-AUG-1999;	PR
11-AUG-1999;	PR
12-AUG-1999;	PR
13-AUG-1999;	PR
13-AUG-1999;	PR
13-AUG-1999;	PR
16-AUG-1999;	PR
17-AUG-1999;	PR
18-AUG-1999;	PR
20-AUG-1999;	PR
20-AUG-1999;	PR
23-AUG-1999;	PR
23-AUG-1999;	PR
23-AUG-1999;	PR
25-AUG-1999;	PR
26-AUG-1999;	PR
27-AUG-1999;	PR
27-AUG-1999;	PR
27-AUG-1999;	PR
30-AUG-1999;	PR
31-AUG-1999;	PR
01-SEP-1999;	PR
07-SEP-1999;	PR
10-SEP-1999;	PR
13-SEP-1999;	PR
15-SEP-1999;	PR
16-SEP-1999;	PR
20-SEP-1999;	PR
22-SEP-1999;	PR
23-SEP-1999;	PR
24-SEP-1999;	PR
28-SEP-1999;	PR
29-SEP-1999;	PR
04-OCT-1999;	PR
05-OCT-1999;	PR
06-OCT-1999;	PR
07-OCT-1999;	PR
08-OCT-1999;	PR
12-OCT-1999;	PR
13-OCT-1999;	PR
13-OCT-1999;	PR
13-OCT-1999;	PR
14-OCT-1999;	PR
14-OCT-1999;	PR
14-OCT-1999;	PR
14-OCT-1999;	PR
18-OCT-1999;	PR
18-OCT-1999;	PR
21-OCT-1999;	PR
21-OCT-1999;	PR
21-OCT-1999;	PR
22-OCT-1999;	PR
22-OCT-1999;	PR
22-OCT-1999;	PR
25-OCT-1999;	PR
25-OCT-1999;	PR
25-OCT-1999;	PR
26-OCT-1999;	PR
26-OCT-1999;	PR
26-OCT-1999;	PR
28-OCT-1999;	PR
28-OCT-1999;	PR
29-OCT-1999;	PR

```

Query Match      34.2%; Score 95.5.; DB3; Length 44;
Best Local Similarity 42.3%; Ref. No. 1.1e-79;
Matches 181; Conservative 102; Mismatches 140; Indels 5; Gaps 2;
100 FLCTLI VALGPIGFTGTSFSSPTDAMVRDNLNLSIFPSAFGLSNVGMVGAIAAGQM 159

```

Db 13 YLSTFVAVCGPFAFGSCAGYSSPAQAARNDLSLTIAEFSLFGSLLTFGAMIGAITSGPI 72 PR 06-MAY-1999; 99US-0132487P.
Qy 160 AEYGRKSLMAIPNIIGHWLAISPAKDAFLYNGRLLEGFGVGIISYVYVVIABISP 219 PR 07-MAY-1999; 99US-0132863P.
Db 73 ADLVGRKGMARVSFAFCVVGWLAIFPAKGVVVALDLGLATGYGMAFSYVYPIFABIAIP 132 PR 11-MAY-1999; 99US-0134256P.
Qy 220 QNMRGALSVNQLSVTFGIFLAYLLGNFIPWRLMVAICAPCTMLIPCLFFIPESPRWLA 279 PR 14-MAY-1999; 99US-0134219P.
Db 133 KTFRGALTTINQLICTGVSVSFIIGTLVTVRWVLAIGIPCAASFLGLFFIPESPRWLA 192 PR 14-MAY-1999; 99US-0134221P.
Qy 280 KMNLTEDCETSLQVLRPETDITTEVDNDIKRANASSKRTTISFQELNQKXYRTPLLGI 339 PR 18-MAY-1999; 99US-0134370P.
Db 193 KVGSDTEFEALRLKRGKADISEEAAEIODYIETLERLPKAKMLDLFQRRYIRSVLIAF 252 PR 18-MAY-1999; 99US-0134768P.
Qy 340 GLLVLQNLGSLNGVLFYASSIFKAAVTNSDLATCSLGAIOVLATGYTWTLLDRAGRIL 399 PR 19-MAY-1999; 99US-0134941P.
Db 253 GLMVFQFGGNGICFTSSIFEQAGPP-TRLGMIIVAVLQVITANALNAVIVDRAGKPL 311 PR 20-MAY-1999; 99US-0135124P.
Qy 400 LIISTGNTLCLLAVSVVFFLKNISQDSNSYIILTMISLVGIVSFVITFSFGNGAIPWL 459 PR 21-MAY-1999; 99US-0135353P.
Db 312 LLVSATGLVIGCLIAAVSVFLKVH---DMAHEAVPVLAVVGIMVYIGSSAGMGAMPWV 367 PR 24-MAY-1999; 99US-0135629P.
Qy 460 MYSILPVSILSGSSTATLANWLTSFAITMTNMLTWSVGSTFSLSYMVVSATFIVFV 519 PR 25-MAY-1999; 99US-0136021P.
Db 368 VMSIFFINIKGVAGGNATLVNFWGAWASYTFNFLMSWSYSGTFLIYAANALAIIVFVI 427 PR 27-MAY-1999; 99US-0136392P.
Qy 520 LWPETKG 527 PR 28-MAY-1999; 99US-0136782P.
Db 428 AIVPETKG 435 PR 01-JUN-1999; 99US-0137222P.
RESULT 5 PR 03-JUN-1999; 99US-0137528P.
AAG13748 AAG13748 standard; protein; 463 AA. PR 04-JUN-1999; 99US-0137502P.
AC AAG13748; PR 07-JUN-1999; 99US-0137724P.
DT 17-OCT-2000 (first entry) PR 08-JUN-1999; 99US-0138094P.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 13357. PR 10-JUN-1999; 99US-0138540P.
KW Protein identification; signal transduction pathway; metabolic pathway; PR 10-JUN-1999; 99US-0138847P.
KW hybridisation assay; genetic mapping; gene expression control; promoter; PR 14-JUN-1999; 99US-0139119P.
KW termination sequence. PR 16-JUN-1999; 99US-0139452P.
OS Arabidopsis thaliana. PR 16-JUN-1999; 99US-0139453P.
XX EP1033405-A2. PR 17-JUN-1999; 99US-0139492P.
XX 06-SEP-2000. PR 18-JUN-1999; 99US-0139454P.
XX 25-FEB-2000; 2000EP-00301439. PR 18-JUN-1999; 99US-0139455P.
XX 25-FEB-1999; 99US-0121825P. PR 18-JUN-1999; 99US-0139461P.
PR 05-MAR-1999; 99US-01233180P. PR 18-JUN-1999; 99US-0139463P.
PR 09-MAR-1999; 99US-01233548P. PR 21-JUN-1999; 99US-0139817P.
PR 23-MAR-1999; 99US-0125788P. PR 22-JUN-1999; 99US-0139899P.
PR 25-MAR-1999; 99US-0126264P. PR 23-JUN-1999; 99US-0140353P.
PR 29-MAR-1999; 99US-0126785P. PR 24-JUN-1999; 99US-0140635P.
PR 01-APR-1999; 99US-0127462P. PR 28-JUN-1999; 99US-0140823P.
PR 06-APR-1999; 99US-0128234P. PR 29-JUN-1999; 99US-0140991P.
PR 08-APR-1999; 99US-0128714P. PR 30-JUN-1999; 99US-0141287P.
PR 16-APR-1999; 99US-0129845P. PR 01-JUL-1999; 99US-0141842P.
PR 19-APR-1999; 99US-0130077P. PR 02-JUL-1999; 99US-0142055P.
PR 21-APR-1999; 99US-0130449P. PR 06-JUL-1999; 99US-0142330P.
PR 23-APR-1999; 99US-0130510P. PR 08-JUL-1999; 99US-0142803P.
PR 28-APR-1999; 99US-0130891P. PR 09-JUL-1999; 99US-0142920P.
PR 30-APR-1999; 99US-0131449P. PR 12-JUL-1999; 99US-0142977P.
PR 30-APR-1999; 99US-0132048P. PR 13-JUL-1999; 99US-0143542P.
PR 04-MAY-1999; 99US-0132407P. PR 14-JUL-1999; 99US-0143624P.
PR 05-MAY-1999; 99US-0132484P. PR 15-JUL-1999; 99US-0144005P.
PR 06-MAY-1999; 99US-0132485P. PR 16-JUL-1999; 99US-0144085P.
PR 99US-0144086P. PR 19-JUL-1999; 99US-0144331P.
PR 99US-0144332P. PR 19-JUL-1999; 99US-0144333P.
PR 99US-0144334P. PR 19-JUL-1999; 99US-0144335P.
PR 99US-0144352P. PR 20-JUL-1999; 99US-0144632P.
PR 99US-0144632P. PR 20-JUL-1999; 99US-0144684P.
PR 99US-0144814P. PR 21-JUL-1999; 99US-0145086P.
PR 99US-0145086P. PR 21-JUL-1999; 99US-0145086P.
PR 99US-0145085P. PR 22-JUL-1999; 99US-0145085P.
PR 99US-0145087P. PR 22-JUL-1999; 99US-0145087P.
PR 99US-0145089P. PR 22-JUL-1999; 99US-0145089P.

PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145313P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145318P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	27-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161381P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161992P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.	Query Match 34.2%; Score 951.5; DB 3; Length 463;		
PR	05-AUG-1999;	99US-0147192P.	Best Local Similarity 42.3%; Pred. No. 1.2e-79;		
PR	05-AUG-1999;	99US-0147260P.	Matches 181; Conservative 102; Mismatches 140; Indels 5; Gaps 2;		
PR	06-AUG-1999;	99US-0147303P.	QY	100	FLCTLVALGPIQFQFTSGFSPQDAMVRDLNTSISFSAFGSLNSVGVNGVIAISQOM 159
PR	06-AUG-1999;	99US-0147416P.	DB	29	YLSTFVAVCGPPAFGSCAGYSSPAQAARNDLSLTIAEFSLGSLLTGAMIGAITSGPI 88
PR	08-AUG-1999;	99US-0147492P.	QY	160	ABYIGRKSMLIAIPNIIIGHIAISFAKXDAFLYMGRLLEGVGLISYTPVVIATISP 219
PR	09-AUG-1999;	99US-0147935P.	DB	89	ADLVGRKGMARVSSAFVCVVGWLAIFAKGVVALDGLRLATGTGMGAFSYVVPFIAEIAF 148
PR	10-AUG-1999;	99US-0148171P.	QY	220	QNMRCALGSVNQSLVTFIGIFLAYLLGMFIPWRLLAVIGALPCTMLIPGLFFIPESPRMLA 279
PR	11-AUG-1999;	99US-0148319P.	DB	149	KTFRCALTILAQILICTGVSVSFIIGTLVTRVLALIGIIPCAASFLGLFFIPESPRMLA 208
PR	13-AUG-1999;	99US-0148565P.	QY	280	KVNLTEDCETSLQVLRGPETDITTEVNDIKRAVASSKRTTISFOELNQKVKYETPLLGI 339
PR	13-AUG-1999;	99US-0148684P.	DB	209	KVGRDTEFEALRKLGRKKADI SEBAEIQDIETLERLPKAKMDLDFQRRVIRSVLIAF 268
PR	16-AUG-1999;	99US-0149358P.	QY	340	GLLVQLNLSGINGVLFYASSIFKAAAGVTNSDLATCSLGAIOVLATGVTTLWLLDRAGRIIL 399
PR	17-AUG-1999;	99US-0149175P.	DB	269	GLMVFQQFGINGICFYITSSIFEQAGFP-TRLGMIYAVLVQVITLALNAPIVDRAGRPEL 327
PR	18-AUG-1999;	99US-0149426P.	QY	400	LIISTSGMTCLLAVSVVFFLKNISQDSNSVYIITMISLVGIVSFVITFSFGMCAIPWL 459
PR	20-AUG-1999;	99US-0149723P.	DB	328	LLVSATGLVIGCLIAAVSFYLVH---DMAHEAVPVLAVGVIMVYIGSFSGMGMAMPWV 383
PR	20-AUG-1999;	99US-0149929P.	QY	460	MMSEILPVSIKSLGSIATLANLTPFAITMTTNLMTWSVGGTFLSYMVVSAFTIVFV 519
PR	23-AUG-1999;	99US-0149902P.	DB	384	VNSEIPPINIKGVAGGMATLVNFWGAWAVSYTFNFMSSVYGTFLIYAINALAIVFVI 443
PR	23-AUG-1999;	99US-0149930P.	QY	520	LMVPETKG 527
PR	25-AUG-1999;	99US-0150566P.	DB	444	AIVPETKG 451
PR	26-AUG-1999;	99US-0150884P.	RESULT 6		
PR	27-AUG-1999;	99US-0151066P.	AAG20978		
PR	30-AUG-1999;	99US-0151303P.	ID AAG20978 standard; protein; 437 AA.		
PR	31-AUG-1999;	99US-0151438P.	XX	AC	AAG20978;
PR	01-SEP-1999;	99US-0151930P.	XX	DT	17-OCT-2000 (first entry)
PR	07-SEP-1999;	99US-0152363P.	XX	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 23367.
PR	10-SEP-1999;	99US-0153070P.	XX	XX	Protein identification; signal transduction pathway; metabolic pathway;
PR	13-SEP-1999;	99US-0153758P.	XX	XX	Hybridisation assay; genetic mapping; gene expression control; promoter;
PR	15-SEP-1999;	99US-0154018P.	XX	XX	termination sequence.
PR	16-SEP-1999;	99US-0154039P.	XX	OS	Arabidopsis thaliana.
PR	20-SEP-1999;	99US-0154779P.	XX	XX	Arabidopsis thaliana.
PR	22-SEP-1999;	99US-0155139P.	XX	XX	EPI033405-A2.
PR	23-SEP-1999;	99US-0155486P.	XX	XX	06-SEP-2000.
PR	24-SEP-1999;	99US-0155659P.	XX	XX	25-FEB-2000; 2000EP-00301439.
PR	28-SEP-1999;	99US-0156458P.	XX	XX	
PR	29-SEP-1999;	99US-0156596P.	XX	XX	
PR	04-OCT-1999;	99US-0157117P.	XX	XX	
PR	05-OCT-1999;	99US-0157533P.	XX	XX	
PR	06-OCT-1999;	99US-0157865P.	XX	XX	
PR	07-OCT-1999;	99US-0158029P.	XX	XX	
PR	08-OCT-1999;	99US-0158232P.	XX	XX	
PR	12-OCT-1999;	99US-0158369P.	XX	XX	
PR	13-OCT-1999;	99US-0159293P.	XX	XX	
PR	13-OCT-1999;	99US-0159294P.	XX	XX	
PR	13-OCT-1999;	99US-0159295P.	XX	XX	
PR	14-OCT-1999;	99US-0159322P.	XX	XX	
PR	14-OCT-1999;	99US-0159330P.	XX	XX	
PR	14-OCT-1999;	99US-0159331P.	XX	XX	
PR	14-OCT-1999;	99US-0159332P.	XX	XX	
PR	14-OCT-1999;	99US-0159637P.	XX	XX	
PR	14-OCT-1999;	99US-0159638P.	XX	XX	
PR	18-OCT-1999;	99US-0159584P.	XX	XX	
PR	21-OCT-1999;	99US-0160741P.	XX	XX	
PR	21-OCT-1999;	99US-0160767P.	XX	XX	
PR	21-OCT-1999;	99US-0160768P.	XX	XX	
PR	21-OCT-1999;	99US-0160770P.	XX	XX	
PR	21-OCT-1999;	99US-0160814P.	XX	XX	

XX	PR	25-FEB-1999;	99US-0121825P.	PR	12-JUL-1999;	99US-0142977P.
PR	PR	05-MAR-1999;	99US-0123180P.	PR	13-JUL-1999;	99US-0143542P.
PR	PR	09-MAR-1999;	99US-0123548P.	PR	14-JUL-1999;	99US-0143624P.
PR	PR	23-MAR-1999;	99US-0123789P.	PR	15-JUL-1999;	99US-0144005P.
PR	PR	23-MAR-1999;	99US-0126264P.	PR	16-JUL-1999;	99US-0144085P.
PR	PR	29-MAR-1999;	99US-0126785P.	PR	16-JUL-1999;	99US-0144086P.
PR	PR	01-APR-1999;	99US-0127462P.	PR	19-JUL-1999;	99US-0144325P.
PR	PR	06-APR-1999;	99US-0128234P.	PR	19-JUL-1999;	99US-0144331P.
PR	PR	08-APR-1999;	99US-0128714P.	PR	19-JUL-1999;	99US-0144334P.
PR	PR	16-APR-1999;	99US-0129845P.	PR	19-JUL-1999;	99US-0144333P.
PR	PR	19-APR-1999;	99US-0130077P.	PR	19-JUL-1999;	99US-0144334P.
PR	PR	21-APR-1999;	99US-0130449P.	PR	19-JUL-1999;	99US-0144335P.
PR	PR	23-APR-1999;	99US-0130510P.	PR	20-JUL-1999;	99US-0144352P.
PR	PR	23-APR-1999;	99US-0130891P.	PR	20-JUL-1999;	99US-0144632P.
PR	PR	28-APR-1999;	99US-0131449P.	PR	20-JUL-1999;	99US-0144684P.
PR	PR	30-APR-1999;	99US-0132048P.	PR	21-JUL-1999;	99US-0144814P.
PR	PR	04-MAY-1999;	99US-0132407P.	PR	21-JUL-1999;	99US-0145086P.
PR	PR	05-MAY-1999;	99US-0132484P.	PR	21-JUL-1999;	99US-0145088P.
PR	PR	06-MAY-1999;	99US-0132486P.	PR	22-JUL-1999;	99US-0145085P.
PR	PR	06-MAY-1999;	99US-0132863P.	PR	22-JUL-1999;	99US-0145087P.
PR	PR	07-MAY-1999;	99US-0132863P.	PR	22-JUL-1999;	99US-0145089P.
PR	PR	11-MAY-1999;	99US-0134256P.	PR	22-JUL-1999;	99US-0145192P.
PR	PR	14-MAY-1999;	99US-0134218P.	PR	23-JUL-1999;	99US-0145145P.
PR	PR	14-MAY-1999;	99US-0134219P.	PR	23-JUL-1999;	99US-0145221P.
PR	PR	14-MAY-1999;	99US-0134221P.	PR	26-JUL-1999;	99US-0145276P.
PR	PR	14-MAY-1999;	99US-0134370P.	PR	27-JUL-1999;	99US-0145913P.
PR	PR	18-MAY-1999;	99US-0134768P.	PR	27-JUL-1999;	99US-0145918P.
PR	PR	19-MAY-1999;	99US-0134941P.	PR	27-JUL-1999;	99US-0145919P.
PR	PR	20-MAY-1999;	99US-0135124P.	PR	28-JUL-1999;	99US-0145951P.
PR	PR	21-MAY-1999;	99US-0135353P.	PR	02-AUG-1999;	99US-0146386P.
PR	PR	24-MAY-1999;	99US-0135629P.	PR	02-AUG-1999;	99US-0146388P.
PR	PR	25-MAY-1999;	99US-0136021P.	PR	02-AUG-1999;	99US-0146389P.
PR	PR	27-MAY-1999;	99US-0136392P.	PR	03-AUG-1999;	99US-0147038P.
PR	PR	28-MAY-1999;	99US-0136782P.	PR	04-AUG-1999;	99US-0147204P.
PR	PR	01-JUN-1999;	99US-0137222P.	PR	04-AUG-1999;	99US-0147302P.
PR	PR	03-JUN-1999;	99US-0137528P.	PR	05-AUG-1999;	99US-0147182P.
PR	PR	04-JUN-1999;	99US-0137502P.	PR	05-AUG-1999;	99US-0147260P.
PR	PR	07-JUN-1999;	99US-0137724P.	PR	06-AUG-1999;	99US-0147303P.
PR	PR	08-JUN-1999;	99US-0138094P.	PR	06-AUG-1999;	99US-0147416P.
PR	PR	10-JUN-1999;	99US-0138540P.	PR	09-AUG-1999;	99US-0147493P.
PR	PR	10-JUN-1999;	99US-0138847P.	PR	09-AUG-1999;	99US-0147935P.
PR	PR	14-JUN-1999;	99US-0139119P.	PR	10-AUG-1999;	99US-0148171P.
PR	PR	16-JUN-1999;	99US-0139452P.	PR	11-AUG-1999;	99US-0148319P.
PR	PR	16-JUN-1999;	99US-0139453P.	PR	12-AUG-1999;	99US-0148341P.
PR	PR	17-JUN-1999;	99US-0139492P.	PR	13-AUG-1999;	99US-0148565P.
PR	PR	18-JUN-1999;	99US-0139454P.	PR	13-AUG-1999;	99US-0148684P.
PR	PR	18-JUN-1999;	99US-0139455P.	PR	16-AUG-1999;	99US-0149368P.
PR	PR	18-JUN-1999;	99US-0139456P.	PR	17-AUG-1999;	99US-0149175P.
PR	PR	18-JUN-1999;	99US-0139457P.	PR	18-AUG-1999;	99US-0149426P.
PR	PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149722P.
PR	PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149723P.
PR	PR	18-JUN-1999;	99US-0139460P.	PR	20-AUG-1999;	99US-0149929P.
PR	PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149902P.
PR	PR	18-JUN-1999;	99US-0139462P.	PR	23-AUG-1999;	99US-0149930P.
PR	PR	18-JUN-1999;	99US-0139463P.	PR	25-AUG-1999;	99US-0150566P.
PR	PR	18-JUN-1999;	99US-0139750P.	PR	26-AUG-1999;	99US-0150884P.
PR	PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151065P.
PR	PR	22-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151066P.
PR	PR	22-JUN-1999;	99US-0139899P.	PR	27-AUG-1999;	99US-0151066P.
PR	PR	23-JUN-1999;	99US-0140353P.	PR	30-AUG-1999;	99US-0151303P.
PR	PR	23-JUN-1999;	99US-0140354P.	PR	31-AUG-1999;	99US-0151438P.
PR	PR	24-JUN-1999;	99US-0140695P.	PR	01-SEP-1999;	99US-0151930P.
PR	PR	24-JUN-1999;	99US-0140823P.	PR	07-SEP-1999;	99US-0152363P.
PR	PR	28-JUN-1999;	99US-0140921P.	PR	10-SEP-1999;	99US-0153070P.
PR	PR	30-JUN-1999;	99US-0141287P.	PR	13-SEP-1999;	99US-0153758P.
PR	PR	01-JUL-1999;	99US-0141842P.	PR	13-SEP-1999;	99US-0154018P.
PR	PR	01-JUL-1999;	99US-0142154P.	PR	16-SEP-1999;	99US-0154039P.
PR	PR	02-JUL-1999;	99US-0142055P.	PR	20-SEP-1999;	99US-0154779P.
PR	PR	06-JUL-1999;	99US-0142390P.	PR	22-SEP-1999;	99US-0155139P.
PR	PR	08-JUL-1999;	99US-0142803P.	PR	23-SEP-1999;	99US-0155486P.
PR	PR	09-JUL-1999;	99US-0142920P.	PR	24-SEP-1999;	99US-0155659P.
PR	PR			PR	28-SEP-1999;	99US-0156458P.

PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159638P.
PR	21-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	22-OCT-1999;	99US-0161404P.
PR	23-OCT-1999;	99US-0161406P.
PR	25-OCT-1999;	99US-0161406P.
PR	25-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161320P.
PR	28-OCT-1999;	99US-0161392P.
PR	28-OCT-1999;	99US-0161393P.
PR	28-OCT-1999;	99US-0162142P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 34.0%; Score 943.5; DB 3; Length 437;		
Best Local Similarity 42.1%; Pred. No. 5.9e-79;		
Matches 180; Conservative 102; Mismatches 141; Indels 5; Gaps 2;		
QY	100	FLCTLVALGPIQFGTSGSSSTQDAMVDDNLISERSAFGSLNMGVGCALASQM 159
Db	3	YLSTFVAVCGSFAGSCAGYSPAQAIRNDLSLTIAEFSLFGSLTFCGAMIGAITSGPI 62
QY	160	AEYIGRGKSLMIAIPNIIGWLAIISPAKDAFLYMGRLLEGFGVGIISYVTVYIAEIGP 219
Db	63	ADLVGRKAMRVSSAFVCGWLAIIFAKGVVVDLGLGELATGYGMAFSYVVPFIAEIAIP 122
QY	220	QNMKGALGSYNQLSVTFEIGFAYLLGMFIPIWRLAVIGALPCTMLIPGLFFIPESRWLA 279
Db	123	KTRFGALTLNQILICGVSVSFIQTLVTRVRLALIGIIPCAASFGLGFFIPESRWLA 182
QY	280	KMNLTEPCETSLQVLRGFETDITTEVNDIKRAVASSKRTTISFQELNOKKYRTPLLLGI 339
Db	183	KVGEDTEFAALRKIRGKKADISEAAEIOQYIETLERLPKAKMLDLFQRYIERSVLIAF 242
QY	340	GLLVQLNLSGNGVLYFVASSIFKAAAGTNSDLATCSLGAIQVLATGVTTWLLDRAGRRIL 399
Db	243	GLMWYFQGGINGICFTYSIFEQAGFP-IRLGMIIYAVLQVITALNAPIVDRAGRKPL 301
QY	400	LIISTGNTCLLAVSVVFFLKNISODSNSYILTMISLVGLVSVFVITFSFGMAIPWL 459
Db	302	LLVATGLVIGCLIAAASVFLKVH-DMHAHEAVPVAVGVVIGVSGFSGAGMPVW 357
QY	460	MMSEILPVSKISGLGSIATLANWLTSFAITWTNMLTWSVGGTFLSMVVSFAFTVFV 519
Db	358	VMSEIFPINIKGVAGGNATLVNWFAGAVSVTFNFMNSWSYGTFLIYAINALIAVFVI 417
QY	520	LWVPETKG 527
Db	418	AIVPETKG 425

RESULT 7
AAG20977
ID AAG20977 standard; protein; 447 AA.
XX
AC AAG20977;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 23366.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127482P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138548P.
PR 14-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.

PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149939P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	23-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-0151338P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144086P.	PR	07-OCT-1999;	99US-0158039P.
PR	16-JUL-1999;	99US-0144325P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144331P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144484P.	PR	14-OCT-1999;	99US-0159637P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160814P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160981P.
PR	26-JUL-1999;	99US-0145276P.	PR	22-OCT-1999;	99US-0160989P.
PR	27-JUL-1999;	99US-0145313P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145318P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145391P.	PR	25-OCT-1999;	99US-0161406P.
PR	28-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	03-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161992P.
PR	04-AUG-1999;	99US-0147204P.	PR	28-OCT-1999;	99US-0161993P.
PR	05-AUG-1999;	99US-0147302P.	PR	29-OCT-1999;	99US-0162142P.
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			

Query Match 34.0%; Score 943.5; DB 3; Length 447;
 Best Local Similarity 42.1%; Pred. No. 6.1e-79;
 Matches 180; Conservative 102; Mismatches 141; Indels 5; Gaps 2;

Qy	100	FLCTLVALGPIQFGTSGFSSPTQDAMVRDLNLSISEFSAFGSLNVMGMAIASGM	159
Db	13	YLSFVAVCGSPAFGSCAGYSSPAQAIRNDLSLTIAEFSLFGSLTFTGAMIGAITSGPI	72
Qy	160	AEYIGRKGLMIAIPIINICMLAISAKDASFLYNGRLLEGFCGVGIISYTVVYIAEISP	219
Db	73	ADLVGRKGMRVSAFCVGVGLAIIFAKGVVADLGLRLATGYGMGAFSVVPIFAEIAIP	132
Qy	220	QNMREGALGSYNQLSVTFGIFLAYLGNFIPWRLLAVIGALPCTMLIPGLFFIPESPRWLA	279
Db	133	KTRFGALTTINQLICTGVSVSFIIGTLVTRVLAIGIIPCAASFLGLFFIPESPRWLA	192

QY 280 KNNLTDCSTSLQVLGPFETDITTEVNDIKRAVASSKRTTISFOELNOKKYRTPELLGI 339
Db 193 KVGRTTEFAARLKRGKXADISEAAEQDYETIELERPKAKMLDLQRRVIRSVLIAF 252
QY 340 GLVLQNLGSGINGVLFIYASSIFKAAQVNTSDSLATCSLGAIQVATGVTWLLDRAGRRL 399
Db 253 GLMVFOQGGINGICFTYSIFEQAGFP-TRLGMIYAVLQVVITLALNAPIVDRAGRKPL 311
QY 400 LIISTSGMTCLLAVSVFELKXDNISQDSNXYILTMISLGVISGVITTFSGMGAIPWL 459
Db 312 LNVSATGLVIGCLIAAVSYFLKH-----DMAHEAVPLVAVGIMVYIGSFSAGMGMPV 367
QY 460 MMSEILPVSIKSGSGLATLANWLTSAITMTTNMLTWSVGGTFLSNMVVSAFTIVFV 519
Db 368 VMSEIFPNIKGWAGGMATLVNWFAGAVSYTFNFMLSMSSYGTFLIYAALAIIVFVI 427
QY 520 LWPVETKG 527
Db 428 AIVPETKG 435
RESULT 8
AAG20976
ID AAG20976 standard; protein; 463 AA.
XX AC AAG20976;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 23365.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145916P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.

PR	16-APR-1999;	99US-0129845P.	PR	19-JUL-1999;	99US-0144334P.
PR	19-APR-1999;	99US-0130077P.	PR	19-JUL-1999;	99US-0144335P.
PR	21-APR-1999;	99US-0130449P.	PR	20-JUL-1999;	99US-0144352P.
PR	23-APR-1999;	99US-0130510P.	PR	20-JUL-1999;	99US-0144632P.
PR	23-APR-1999;	99US-0130891P.	PR	20-JUL-1999;	99US-0144884P.
PR	28-APR-1999;	99US-0131442P.	PR	21-JUL-1999;	99US-0144814P.
PR	30-APR-1999;	99US-0132048P.	PR	21-JUL-1999;	99US-0145086P.
PR	30-APR-1999;	99US-0132407P.	PR	21-JUL-1999;	99US-0145088P.
PR	04-MAY-1999;	99US-0132484P.	PR	22-JUL-1999;	99US-0145087P.
PR	05-MAY-1999;	99US-0132485P.	PR	22-JUL-1999;	99US-0145089P.
PR	06-MAY-1999;	99US-0132486P.	PR	22-JUL-1999;	99US-0145192P.
PR	06-MAY-1999;	99US-0132487P.	PR	23-JUL-1999;	99US-0145145P.
PR	07-MAY-1999;	99US-0132863P.	PR	23-JUL-1999;	99US-0145218P.
PR	11-MAY-1999;	99US-0134256P.	PR	23-JUL-1999;	99US-0145224P.
PR	14-MAY-1999;	99US-0134218P.	PR	26-JUL-1999;	99US-0145276P.
PR	14-MAY-1999;	99US-0134219P.	PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999;	99US-0134221P.	PR	27-JUL-1999;	99US-0145918P.
PR	14-MAY-1999;	99US-0134370P.	PR	27-JUL-1999;	99US-0145919P.
PR	18-MAY-1999;	99US-0134768P.	PR	28-JUL-1999;	99US-0145951P.
PR	19-MAY-1999;	99US-0134941P.	PR	02-AUG-1999;	99US-0146386P.
PR	20-MAY-1999;	99US-0135124P.	PR	02-AUG-1999;	99US-0146388P.
PR	21-MAY-1999;	99US-0135629P.	PR	02-AUG-1999;	99US-0146389P.
PR	24-MAY-1999;	99US-0136021P.	PR	03-AUG-1999;	99US-0147038P.
PR	25-MAY-1999;	99US-0136392P.	PR	04-AUG-1999;	99US-0147204P.
PR	27-MAY-1999;	99US-0136782P.	PR	04-AUG-1999;	99US-0147302P.
PR	28-MAY-1999;	99US-0137222P.	PR	05-AUG-1999;	99US-0147192P.
PR	01-JUN-1999;	99US-0137528P.	PR	05-AUG-1999;	99US-0147260P.
PR	03-JUN-1999;	99US-0137528P.	PR	06-AUG-1999;	99US-0147303P.
PR	04-JUN-1999;	99US-0137502P.	PR	06-AUG-1999;	99US-0147416P.
PR	07-JUN-1999;	99US-0137724P.	PR	09-AUG-1999;	99US-0147493P.
PR	08-JUN-1999;	99US-0138094P.	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	99US-0138540P.	PR	10-AUG-1999;	99US-0148171P.
PR	10-JUN-1999;	99US-0138847P.	PR	11-AUG-1999;	99US-0148319P.
PR	14-JUN-1999;	99US-0139119P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-0139452P.	PR	13-AUG-1999;	99US-0148555P.
PR	16-JUN-1999;	99US-0139453P.	PR	13-AUG-1999;	99US-0148556P.
PR	17-JUN-1999;	99US-0139453P.	PR	16-AUG-1999;	99US-0149388P.
PR	18-JUN-1999;	99US-0139454P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139455P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139456P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139457P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149939P.
PR	18-JUN-1999;	99US-0139459P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139462P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139463P.	PR	27-AUG-1999;	99US-0151066P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151080P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151303P.
PR	21-JUN-1999;	99US-0139817P.	PR	30-AUG-1999;	99US-0151303P.
PR	21-JUN-1999;	99US-0139899P.	PR	31-AUG-1999;	99US-0151438P.
PR	22-JUN-1999;	99US-0140353P.	PR	01-SEP-1999;	99US-0151930P.
PR	23-JUN-1999;	99US-0140354P.	PR	07-SEP-1999;	99US-0152363P.
PR	23-JUN-1999;	99US-0140354P.	PR	10-SEP-1999;	99US-0153070P.
PR	24-JUN-1999;	99US-0140895P.	PR	13-SEP-1999;	99US-0153758P.
PR	28-JUN-1999;	99US-0140823P.	PR	13-SEP-1999;	99US-0153758P.
PR	28-JUN-1999;	99US-0140823P.	PR	15-SEP-1999;	99US-0154018P.
PR	29-JUN-1999;	99US-0140951P.	PR	16-SEP-1999;	99US-0154039P.
PR	30-JUN-1999;	99US-0141287P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0141842P.	PR	20-SEP-1999;	99US-0154779P.
PR	01-JUL-1999;	99US-0142154P.	PR	22-SEP-1999;	99US-0155139P.
PR	02-JUL-1999;	99US-0142055P.	PR	23-SEP-1999;	99US-0155486P.
PR	06-JUL-1999;	99US-0142330P.	PR	23-SEP-1999;	99US-0155659P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0156458P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-015753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158323P.
PR	19-JUL-1999;	99US-0144332P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0158293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.

PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.	PR	07-SEP-1999;	99US-0152363P.	PR	10-SEP-1999;	99US-0153070P.	PR	13-SEP-1999;	99US-0153758P.	PR	15-SEP-1999;	99US-0154018P.	PR	16-SEP-1999;	99US-0154039P.	PR	20-SEP-1999;	99US-0154779P.	PR	22-SEP-1999;	99US-0155139P.	PR	23-SEP-1999;	99US-0155486P.	PR	24-SEP-1999;	99US-0155659P.	PR	28-SEP-1999;	99US-0156458P.	PR	29-SEP-1999;	99US-0156596P.	PR	04-OCT-1999;	99US-0157117P.	PR	05-OCT-1999;	99US-0157753P.	PR	06-OCT-1999;	99US-0157865P.	PR	07-OCT-1999;	99US-0158029P.	PR	08-OCT-1999;	99US-0158232P.	PR	12-OCT-1999;	99US-0158369P.	PR	13-OCT-1999;	99US-0159293P.	PR	13-OCT-1999;	99US-0159294P.	PR	14-OCT-1999;	99US-0159295P.	PR	14-OCT-1999;	99US-0159339P.	PR	14-OCT-1999;	99US-0159330P.	PR	14-OCT-1999;	99US-0159331P.	PR	14-OCT-1999;	99US-0159637P.	PR	14-OCT-1999;	99US-0159638P.	PR	18-OCT-1999;	99US-0159584P.	PR	21-OCT-1999;	99US-0160741P.	PR	21-OCT-1999;	99US-0160767P.	PR	21-OCT-1999;	99US-0160768P.	PR	21-OCT-1999;	99US-0160770P.	PR	21-OCT-1999;	99US-0160814P.	PR	21-OCT-1999;	99US-0160815P.	PR	22-OCT-1999;	99US-0160980P.	PR	22-OCT-1999;	99US-0160981P.	PR	22-OCT-1999;	99US-0160989P.	PR	23-OCT-1999;	99US-0161404P.	PR	23-OCT-1999;	99US-0161405P.	PR	25-OCT-1999;	99US-0161406P.	PR	25-OCT-1999;	99US-0161539P.	PR	26-OCT-1999;	99US-0161360P.	PR	26-OCT-1999;	99US-0161361P.	PR	28-OCT-1999;	99US-0161920P.	PR	28-OCT-1999;	99US-0161992P.	PR	28-OCT-1999;	99US-0161993P.	PR	29-OCT-1999;	99US-0162142P.	PR	29-OCT-1999;	99US-0162143P.	PR	29-OCT-1999;	99US-0162144P.	PR	29-OCT-1999;	99US-0162145P.	PR	29-OCT-1999;	99US-0162146P.	PR	29-OCT-1999;	99US-0162147P.	PR	29-OCT-1999;	99US-0162148P.	PR	29-OCT-1999;	99US-0162149P.	PR	29-OCT-1999;	99US-0162150P.	PR	29-OCT-1999;	99US-0162151P.	PR	29-OCT-1999;	99US-0162152P.	PR	29-OCT-1999;	99US-0162153P.	PR	29-OCT-1999;	99US-0162154P.	PR	29-OCT-1999;	99US-0162155P.	PR	29-OCT-1999;	99US-0162156P.	PR	29-OCT-1999;	99US-0162157P.	PR	29-OCT-1999;	99US-0162158P.	PR	29-OCT-1999;	99US-0162159P.	PR	29-OCT-1999;	99US-0162160P.	PR	29-OCT-1999;	99US-0162161P.	PR	29-OCT-1999;	99US-0162162P.	PR	29-OCT-1999;	99US-0162163P.	PR	29-OCT-1999;	99US-0162164P.	PR	29-OCT-1999;	99US-0162165P.	PR	29-OCT-1999;	99US-0162166P.	PR	29-OCT-1999;	99US-0162167P.	PR	29-OCT-1999;	99US-0162168P.	PR	29-OCT-1999;	99US-0162169P.	PR	29-OCT-1999;	99US-0162170P.	PR	29-OCT-1999;	99US-0162171P.	PR	29-OCT-1999;	99US-0162172P.	PR	29-OCT-1999;	99US-0162173P.	PR	29-OCT-1999;	99US-0162174P.	PR	29-OCT-1999;	99US-0162175P.	PR	29-OCT-1999;	99US-0162176P.	PR	29-OCT-1999;	99US-0162177P.	PR	29-OCT-1999;	99US-0162178P.	PR	29-OCT-1999;	99US-0162179P.	PR	29-OCT-1999;	99US-0162180P.	PR	29-OCT-1999;	99US-0162181P.	PR	29-OCT-1999;	99US-0162182P.	PR	29-OCT-1999;	99US-0162183P.	PR	29-OCT-1999;	99US-0162184P.	PR	29-OCT-1999;	99US-0162185P.	PR	29-OCT-1999;	99US-0162186P.	PR	29-OCT-1999;	99US-0162187P.	PR	29-OCT-1999;	99US-0162188P.	PR	29-OCT-1999;	99US-0162189P.	PR	29-OCT-1999;	99US-0162190P.	PR	29-OCT-1999;	99US-0162191P.	PR	29-OCT-1999;	99US-0162192P.	PR	29-OCT-1999;	99US-0162193P.	PR	29-OCT-1999;	99US-0162194P.	PR	29-OCT-1999;	99US-0162195P.	PR	29-OCT-1999;	99US-0162196P.	PR	29-OCT-1999;	99US-0162197P.	PR	29-OCT-1999;	99US-0162198P.	PR	29-OCT-1999;	99US-0162199P.	PR	29-OCT-1999;	99US-0162200P.	PR	29-OCT-1999;	99US-0162201P.	PR	29-OCT-1999;	99US-0162202P.	PR	29-OCT-1999;	99US-0162203P.	PR	29-OCT-1999;	99US-0162204P.	PR	29-OCT-1999;	99US-0162205P.	PR	29-OCT-1999;	
----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	----------------	----	--------------	--

Db 312 LLLVSATGLVIGCLIAAVSFYLVKH---DMAHEAVPVLAVVGIMVYIGSFSAGMGAMPW 367
Qy 459 LMKSEILPVKISGLGSIATLANWLTSPALTMTNMLTWSVGGTFLSNVYVSAFTIVFV 518
Db 368 VMSEIFPIKIGVAGGMATLVNMFAGAVSYTFNFMWSYSGTFLIAAINALAIVFV 427
Qy 519 VLVVPTKG 527
Db 428 IAVPTEKG 436

RESULT 11
AAG38869
ID AAG38869 standard; protein; 464 AA.
AC AAG38869;
XX
XX 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 48016.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EPI033405-A2.
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 23-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142330P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.

PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149829P.
PR 22-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 16-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159233P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 33.6%; Score 933; DB 3; Length 464;
Best Local Similarity 42.0%; Pred. No. 6.2e-78;
Matches 180; Conservative 102; Mismatches 141; Indels 6; Gaps 3;
QY 100 FLCTILVALGPIQFGFTSGFSPTQDAMVRDNLNLSIEFSAFGSLSNVGGWGAIASQOM 159

DB 29 YLSTFVAVCGSFAFGSCAGYSSPAQAIRNDLSLTIAEFSLFGSULLTFGAMIGAITSGPI 88
QY 160 AEYIGRKGLMIAAPNIIGWLAIKFAKADASFLYGRLLLEGFGVGIISYTPVPIAISP 219
DB 89 ADLVGRKGAMRVSAFACVVGWLAIFAKGVVALDLGRLATGCGAFSVVPIPIAIAIP 148
QY 220 QNREGALGSVN-QLSVTFGIFLAYLLGMFIPWRLLAVIGALPCTMLIPGLFFIPESPRWL 278
DB 149 KTFRGALTTLNQOILICTGVSVSFIIGFLVTRVALIGIIPCAASFLGLFFIPESPRWL 208
QY 279 AKMNLTEDETSLQVLRGFETDITTEVNDIKRAVASSKRTTISFQELNQKXYTPLL 338
DB 209 AKVGRDTEFEAALRGKRGKADISEAAEQDIYETLERLPKAKMLDLFQRRYIRSVLIA 268
QY 339 IGLVLQNLGSLNGVLFYASSIFKAAAGVTNSDLATCSLGAIVLATGVTWLLDRAGRI 398
DB 269 FGLAVFOQGGINGICFYTSIFQAGPP-TRLGMIIYAVLQVVTALNADIVDRGRKP 327
QY 399 LLIISTGWTCLLAVSVVFFELKDNISQDSNSYILTMISLVGIVSVFVITFSFGGAIPW 458
DB 328 LLLVSATGLVIGCLIAAVSFYLVKH---DMAHEAVPVLAVVGIMVYIGSFSGANGAMPW 383
QY 459 LMMSEILPVSIKSLGGSITATLANWLTSPAITWTTNMLTWSVGGTFLSYMVVSAFTIVFV 518
DB 384 VVMSEIFFINIKVAGGNATLVNMFAGAWASYTFNFLMSWSYSGTFLIYAAINALAIVFV 443
QY 519 VLVVPETKG 527
DB 444 IAVVPETKG 452
RESULT 12
ADM48307
ID ADM48307 standard; protein; 481 AA.
XX ADM48307;
XX
DT 03-JUN-2004 (first entry)
XX
DE Polypeptide sequence #357 useful in producing transgenic plants.
KW Plant; maize; rice; sorghum; thale cress; soybean; wheat; transgenic;
KW osmotic stress; sugar transport; cell cycle pathway; plant height;
KW carbohydrate transport; crop productivity; plant growth;
KW stress resistance; disease resistance; insect resistance; heat tolerance;
KW nitrogen assimilation; water stress tolerance;
KW photosynthetic carbon fixation; virus resistance; gene therapy.
XX Glycine max.
OS
XX
XX US2003233670-A1.
XX
PD 18-DEC-2003.
XX
XX 04-DEC-2002; 2002US-00310154.
XX
XX 04-DEC-2001; 2001US-0337358P.
XX
XX (EDGE/) EDGERTON M D.
XX (CHOM/) CHOMET P S.
XX (LACC/) LACCETTI L B.
XX
XX Edgerton MD, Chomet PS, Laccetti LB;
XX WPI; 2004-061374/06.
XX N-PSDB; ADM47939.
XX
PT New polynucleotide, useful for manipulating plant protein quality,
PT improving plant growth, yield and crop productivity or grain composition
PT or producing plants with improved properties.
XX
XX Claim 8; SEQ ID NO 725; 144bp; English.
PS

XX The present invention relates to polynucleotide sequences, and the
 CC proteins they encode. The sequences are isolated from a variety of
 CC organisms such as plants (e.g. maize, rice, sorghum, thale cress,
 CC soybean, and wheat), cyanobacteria, bacteria, yeast and other fungi. The
 CC polynucleotide and polypeptide sequences of the invention are useful in
 CC the production of transgenic plants that have improved properties. Also
 CC disclosed are methods of producing fertile transgenic plants, preferably
 CC maize, with desired phenotypes. The polynucleotide and polypeptide
 CC sequences are useful for improving plants by providing protection against
 CC osmotic stress, improving altering sugar transport and/or metabolism,
 CC modifying the cell cycle pathway, reducing plant height, modifying
 CC carbohydrate transport, improving crop productivity, improving plant
 CC growth and stress resistance, improving crop productivity, improving plant
 CC insect resistance, improving cold or heat tolerance, improving nitrogen
 CC assimilation, improving stalk strength, improving water stress tolerance,
 CC improving photosynthetic carbon fixation, improving biotic and abiotic
 CC stress resistance, improving resistance to oxidative stress, providing
 CC increased vigour, reducing senescence, and conferring virus resistance.
 CC The present sequence represents a polypeptide sequence of the invention.
 CC Note: The sequence data for this patent is not provided in the printed
 CC specification but is obtained in electronic format from the USPTO website
 CC at seqdata.uspto.gov.

XX Sequence 481 AA;

Query Match 32.1%; Score 891; DB 8; Length 481;

Best Local Similarity 37.2%; Pred. No. 5.4e-74;

Matches 188; Conservative 97; Mismatches 171; Indels 50; Gaps 4;

QY 48 AGAGESSDHDGVLREPLNTGWSYMSRQSPAGTSSMAVLRESHVAELCTLIIVA 107
 DB 25 AAACENGSD-----KSVKNGSIGMVL-----LSTLVAV 53
 QY 108 LGPIQFGTSGFSPTODAMVRDLNLSIFBSAFGLSNVGMVGAIAAGQMAEYIGRGK 167
 DB 54 CGSFTGTCVGSAPTQAARADNLNLSAEFSMFGSLVTIGAMLGAITSGRITDFIRGK 113
 QY 168 SLMTAAIPNIGLWLAISPAKDAFLYNGRLLEGFGVIGIISYTPVVIASIPQNMKGALG 227
 DB 114 AMRISTGFCITGMIAVEFSGVSLDFGRFTTGYGIVSYVVPVVIASIPQNMKGALG 173
 QY 228 SVNQLSVTFGIFLAYLGMFIPWRLAVIAGALPCTMLIPGLFIPSPRWLAKNLTEDC 287
 DB 174 TTNOLLIVTGVSVFLLGVSINRELALAGLVFCILLVGLCFIPSPRWLAKVGEKEF 233
 QY 288 ETSLOVLRGFTDITTEVNDIKRAVASSKRTTISFQELNQKRYTPILLGLLVQLNL 347
 DB 234 QLAISRLRGKADISDEAAEILDYIETQLSPKTKLLDLFQSKYVHSVIVGVGLMACQOS 293
 QY 348 SGINGVLVFIASSIFKAGVNTSLATCSLGAIOVLATGVTMTLLDRAGRIILLISTSM 407
 DB 294 VGINGIGFYFAEIVFVAGLSGKAGTIAVACIQIPFTLIGALMDKSGRPVWVSAAT 353
 QY 408 TLCLLAVSVVFFLKDNIQSDNSYIILTMISLVGIVSVFTFSGMGAIPWLMGSEILPV 467
 DB 354 FLGCFVAAFAFFLKD---QSLPEWVPILAFAGVLIYIAAFSIGLGSVPFWIMSEIPI 409
 QY 468 SIKSLGGSATLANLTSPAITMTMLTWSVGGTFLSMVVSAPTIVFVVLWVPETKG 527
 DB 410 HLKGTAGSLVVLVAVLGMVSVITFNLMSWSPPGLFLYAGCSLLTILFVAKLVPEIKG 469
 QY 528 XNSRGDTIFVSLSIQRLQWLPECLS 553
 DB 470 -----XTLEBIOACIS 480

RESULT 13

AAG39282

ID AAG39282 standard; protein; 468 AA.

XX

AC AAG39282;

XX

DT 18-OCT-2000 (first entry)
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 48581.
 DE
 XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX Arabidopsis thaliana.
 OS Arabidopsis thaliana.
 PN EP1033405-A2.
 XX
 XX 06-SEP-2000.
 XX
 XX 25-FEB-2000; 2000EP-00301439.
 XX
 XX 25-FEB-1999; 99US-0121825P.
 XX 05-MAR-1999; 99US-0123180P.
 XX 09-MAR-1999; 99US-0123488P.
 XX 23-MAR-1999; 99US-0125788P.
 XX 25-MAR-1999; 99US-0126264P.
 XX 29-MAR-1999; 99US-0126785P.
 XX 01-APR-1999; 99US-0127462P.
 XX 06-APR-1999; 99US-0128234P.
 XX 08-APR-1999; 99US-0128714P.
 XX 16-APR-1999; 99US-0129845P.
 XX 19-APR-1999; 99US-0130077P.
 XX 21-APR-1999; 99US-0130449P.
 XX 23-APR-1999; 99US-0130510P.
 XX 28-APR-1999; 99US-0130891P.
 XX 30-APR-1999; 99US-0131449P.
 XX 30-APR-1999; 99US-0132048P.
 XX 04-MAY-1999; 99US-0132407P.
 XX 05-MAY-1999; 99US-0132484P.
 XX 06-MAY-1999; 99US-0132486P.
 XX 06-MAY-1999; 99US-0132487P.
 XX 07-MAY-1999; 99US-0132863P.
 XX 11-MAY-1999; 99US-0134256P.
 XX 14-MAY-1999; 99US-0134218P.
 XX 14-MAY-1999; 99US-0134219P.
 XX 14-MAY-1999; 99US-0134221P.
 XX 14-MAY-1999; 99US-0134370P.
 XX 18-MAY-1999; 99US-0134768P.
 XX 19-MAY-1999; 99US-0134941P.
 XX 20-MAY-1999; 99US-0135124P.
 XX 21-MAY-1999; 99US-0135353P.
 XX 24-MAY-1999; 99US-0135629P.
 XX 25-MAY-1999; 99US-0136021P.
 XX 27-MAY-1999; 99US-0136392P.
 XX 28-MAY-1999; 99US-0136782P.
 XX 01-JUN-1999; 99US-0137222P.
 XX 03-JUN-1999; 99US-0137528P.
 XX 04-JUN-1999; 99US-0137502P.
 XX 07-JUN-1999; 99US-0137724P.
 XX 08-JUN-1999; 99US-0138094P.
 XX 10-JUN-1999; 99US-0138540P.
 XX 10-JUN-1999; 99US-0138847P.
 XX 14-JUN-1999; 99US-0139119P.
 XX 16-JUN-1999; 99US-0139452P.
 XX 16-JUN-1999; 99US-0139453P.
 XX 17-JUN-1999; 99US-0139492P.
 XX 18-JUN-1999; 99US-0139454P.
 XX 18-JUN-1999; 99US-0139455P.
 XX 18-JUN-1999; 99US-0139456P.
 XX 18-JUN-1999; 99US-0139457P.
 XX 18-JUN-1999; 99US-0139458P.
 XX 18-JUN-1999; 99US-0139459P.
 XX 18-JUN-1999; 99US-0139460P.
 XX 18-JUN-1999; 99US-0139461P.
 XX 18-JUN-1999; 99US-0139462P.
 XX 18-JUN-1999; 99US-0139463P.
 XX 18-JUN-1999; 99US-0139750P.


```
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 28-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142603P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 22-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147935P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158023P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 25-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
PR 29-OCT-1999; 99US-0162142P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158023P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 25-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
PR 29-OCT-1999; 99US-0162142P.
Query Match 31.5%; Score 874.5; DB 3; Length 468;
Best Local Similarity 39.4%; Pred. No. 1.8e-72;
Matches 174; Conservative 95; Mismatches 158; Indels 15; Gaps 4;
QY 92 LRESHVSA--FLCTLIVALGPIQFGTSGFSPQDAMVRDLNLSISFSAFGLSNVGG 149
DB 21 INECRITAVLFSTFVSVCGSFCAAGYSSVAQTGINDLGLSVAQYSMFGSMTFGG 80
QY 150 MYGATASGOMAEYIGKGLSLMIAIPNIIGWLAIISFAKDASFLYMGRLLEGFGVGIISYT 209
DB 81 MIGAIFSGKVDLMGKRGKTWFQAQIFCFGWAVAKADSMWLDIGRLSTGFAVGLLSYV 140
QY 210 VPVYIAETSPONRGALGSVNQLSVTFFGIFLAYLGMIPWRLLAIVIGALPCTMLIPGLF 269
DB 141 IPVYIAETPKHVGAFVAFANQMOSGLSLFVYIGNFVHWNALALIGLIPCALQVTLF 200
QY 270 FIPESPRWLAKMNLATEDCETSLQVLRGPETDITTEVNDIKRAVASSSKRTTISFQELNOK 329
DB 201 FIPESPRLLGKMGHEKECRASLQSLRGDDADISEANITKETMILFDEGPKSRVMDLQQR 260
QY 330 KYRTPLLLGIGLLVQLNLSGINGVLFIYASSIFKAAQVNTNSDLATCSLGAIQVLTGVTW 389
```


Query Match 31.5%: Score 874.5: DB 3: Length 470:

PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136332P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140895P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 05-JUL-1999; 99US-0142055P.
PR 05-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147152P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149923P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159233P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.

PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	22-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 31.3%; Score 868.5; DB 3; Length 474;		
Best Local Similarity 41.1%; Pred. No. 6.6e-72;		
Matches 181; Conservative 94; Mismatches 150; Indels 15; Gaps 6;		
Qy	94	ESHVSA--FLCTILVALGPTQGFSGFSPTQDAMVRDNLNLSISFSAFGLSNVGGMV 151
Db	24	ECRITACVILSTFIACGFSFGVSGYTSAGIEMKDDLSIAQFSAPLSILGAAI 83
Qy	152	GATASQMAEYIGRKGSIMIAIENIIGMLAISFAKADASFLYNGRLLEGVGVIISYTP 211
Db	84	GALFSGKMAILGRRTMWTWYDLCLIGWFSIAFAKDVMMWLNFRITSSGIGLGLISYVVP 143
Qy	212	VYIAEISQNMKGALGSVNLSTVFIFLAYLLGMFIPWELLAVIGALPCTMLIPGLFEI 271
Db	144	VYIAEISPKHVRGTFTFTNQLNQLNSGLAMVYFSGNLFNWRILALLGALPCFIQVIGLFFV 203
Qy	272	PESPRMLAKNLTEDCETSQVLRGPETDITTEVNDIK--RAVASSSKRTTISFOELNQ 328
Db	204	PESPRMLAKVSDKELENSLLRLRGNGADISREASDIEVMTKXVENDSKS---SFCDLFQ 260
Qy	329	KYRTPLLLGIGLLVLQNLGNGVLFYASSIFKAAGVTNSDLATCSLGAIQVLATGVTT 388
Db	261	RKYRYTLVWIGLMLIQOFSGSSAVLSYASTILRKAGFSVT-IGSTLLGLFMIPKAMIGV 319
Qy	389	WLDRAGRRLLIISTSGMTLCLLAVSVVFFL-KDNISODSNSYYILTMTISLVGIVSFVI 447
Db	320	ILVDKRRRRLLTFSVGMCIITMLIGVATLQKMLPELTPTVFTFICVTL-----YIG 374
Qy	448	TFSPGMAIPWLMSEILPVSIKSLGGSIAFLANWLTSPAITMTNMLTWSVGGTFLSY 507
Db	375	TYAIGLGLPWIWIMSEIFPNWIKVTAGSIVTLVSWSSSIVTVAFNPLLEWSTQGTFFVF 434
Qy	508	MVVSATFIVFVLWVPETKG 527
Db	435	GAVGGLALLFWLWVPETKG 454

Search completed: October 13, 2004, 11:38:31
Job time : 60.5861 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 11:33:23 ; Search time 17.1663 Seconds
(without alignments)
2136.387 Million cell updates/sec

Title: US-10-051-909-36

Perfect score: 2779
Sequence: 1 PSSSSFRPAGKKKKKNOG.....TIFVLSIQRLQWLPECLLS 553

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/prodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	607	21.8	488	4	US-10-162-012-46
2	575.5	20.7	524	2	Sequence 46, Appl
3	575.5	20.7	524	3	Sequence 12, Appl
4	572.5	20.6	494	2	Sequence 5, Appl
5	572.5	20.6	494	3	Sequence 5, Appl
6	572.5	20.6	494	3	Sequence 5, Appl
7	568	20.4	493	2	Sequence 10, Appl
8	568	20.4	493	3	Sequence 10, Appl
9	568	20.4	493	3	Sequence 10, Appl
10	557.5	20.1	539	3	Sequence 26, Appl
11	550	19.8	509	2	Sequence 6, Appl
12	550	19.8	509	3	Sequence 6, Appl
13	550	19.8	509	3	Sequence 6, Appl
14	550	19.8	509	3	Sequence 22, Appl
15	542.5	19.5	488	2	Sequence 10, Appl
16	542.5	19.5	488	3	Sequence 10, Appl
17	542.5	19.5	523	3	Sequence 24, Appl
18	537.5	19.0	513	3	Sequence 20, Appl
19	524	18.9	863	4	Sequence 9, Appl
20	522	18.8	529	3	Sequence 28, Appl
21	522	18.8	529	3	Sequence 30, Appl
22	514.5	18.5	476	4	Sequence 11933, A
23	497	17.9	514	4	Sequence 11902, A
24	496.5	17.9	584	2	Sequence 13, Appl
25	496.5	17.9	584	3	Sequence 13, Appl
26	493.5	17.8	534	2	Sequence 4, Appl
27	493.5	17.8	534	3	Sequence 4, Appl

28 493.5 17.8 534 3 US-09-610-417-4 Sequence 4, Appli
29 486.5 17.5 492 2 US-08-355-844-3 Sequence 3, Appli
30 486.5 17.5 492 5 PCT-US95-16126-3 Sequence 3, Appli
31 471 16.9 514 4 US-09-679-686B-22 Sequence 22, Appli
32 470.5 16.9 500 2 US-09-031-392-7 Sequence 7, Appli
33 470.5 16.9 500 3 US-09-299-549-7 Sequence 7, Appli
34 470.5 16.9 500 3 US-09-610-417-7 Sequence 7, Appli
35 470.5 16.9 517 4 US-09-679-686B-18 Sequence 18, Appli
36 469 16.9 518 4 US-09-679-686B-23 Sequence 23, Appli
37 469 16.9 519 4 US-09-679-686B-24 Sequence 24, Appli
38 455 16.4 518 4 US-09-679-686B-16 Sequence 16, Appli
39 453.5 16.3 511 4 US-09-679-686B-12 Sequence 12, Appli
40 448 16.3 502 4 US-09-679-686B-2 Sequence 2, Appli
41 446 16.0 383 2 US-09-031-392-3 Sequence 3, Appli
42 446 16.0 383 3 US-09-299-549-3 Sequence 3, Appli
43 446 16.0 383 3 US-09-610-417-3 Sequence 3, Appli
44 446 16.0 501 4 US-09-489-039A-11731 Sequence 11731, A
45 442 15.9 552 4 US-09-248-796A-20713 Sequence 20713, A

ALIGNMENTS

RESULT 1
US-10-162-012-46
; Sequence 46, Application US/10162012
; Patent No. 6682597
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Rory A.J.
; APPLICANT: Silos-Santiago, Inmaculada
; TITLE OF INVENTION: NOVEL HUMAN ION CHANNEL AND TRANSPORTER FAMILY MEMBERS
; FILE REFERENCE: 10448-190001
; CURRENT APPLICATION NUMBER: US/10/162,012
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 60/209,845
; PRIOR FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: US 09/875,321
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: PCT/US01/18340
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/209,257
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/875,423
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US01/18398
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/209,238
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/875,363
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US01/18247
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/227,068
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 09/928,530
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: PCT/US01/25475
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: US 60/226,770
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/934,421
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26096
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/279,281
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 10/109,029
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: PCT/US02/09728
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/290,288
; PRIOR FILING DATE: 2001-05-11

;; PRIOR APPLICATION NUMBER: US (not assigned)
;; PRIOR FILING DATE: 2002-05-13
;; NUMBER OF SEQ ID NOS: 48
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 46
;; LENGTH: 488
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: consensus sequence
US-10-162-012-46

Query Match 21.8%; Score 607; DB 4; Length 488;
Best Local Similarity 34.8%; Pred. No. 7.1e-50;
Matches 150; Conservative 85; Mismatches 140; Indels 56; Gaps 12;
QY 143 SLSNVGMVGAISQVAEYIGRKSIMIAAIPNIIGWLAIAPAKDAS-----FLYM 194
DB 58 SIFPLGRGLISLFGALGDRGRKKSLLIALVFLVIGALLSGAAGPYTTIGLWAFYLLIV 117
QY 195 GRLLGEGFGVGIISYTPVYIAEISPPQNRGALGSVNOLSVTFGIFLAYLLGMFI----- 248
DB 118 GRVLGLGVGASVLPVYISEAPKALRGALGSLVQLAITIGLVAAIIGLGNKNTND 177
QY 249 -----PWRLLAVIGALPCTNMLIPGLFFIPESPRWLAKONMLTDCETSQVLRGPETDIT 302
DB 178 SALNSWGRIPPLGLQLVALLLIGLLFLPESPRWLVEKGLSEAREVLAKLRGVE-DVD 236
QY 303 TEVNDIK--RAVASSSKRTTISFOEL---NOKKYRTPLLGIGLLVLQNLGSGINGVLV 355
DB 237 QEQEIKAELEATVSEKAKASANGELFRGRTPKVRQRLLMGVMLQAFQQLTGINAIFY 296
QY 356 YASIFKAAGVNSD---LATCSLGAIQVLAQTGT-TWLDRAGRRIILLIISTGTMCLCL 411
DB 297 YSPTRKSVGSVDASLLVTIIVGVNVFTEVALIFLVDREGRPLLLGGAAGMAICF 356
QY 412 L-----AVSVVFFL---KDNISQDSNYYLTMTSLVGVSVFVITFSFGNGAIPWLMWSE 463
DB 357 LILGASIGVALLLNKPKDPSKAAG-----IVAIVFLLFAFPALGNGPIPWILSE 410
QY 464 ILPVSIKSLGGSATLANWLTSPA-----ITWTTNMLTWSVGTFSLSYMVVSAFTIV 516
DB 411 LPTKVRKALATAANLANWLTNFIQFLPYITGAIGLALG---GVVLFVAGLLVLFIL 467
QY 517 FVVLWVPETKG 527
DB 468 FVFFVPEPKG 478

RESULT 2
US-06-928-692-12
; Sequence 12, Application US/08928692
; Patent No. 5958727
; GENERAL INFORMATION:
; APPLICANT: Brody, Howard
; APPLICANT: Yaver, Deborah S.
; APPLICANT: Lamsa, Michael
; APPLICANT: Hansen, Kim
; TITLE OF INVENTION: Methods for Modifying the Production of
; TITLE OF INVENTION: a Polypeptide
; NUMBER OF SEQUENCES: 80
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5958727o No. 5958727disk of No. 5958727th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/928.692
;; FILING DATE: 12-SEPT-1997
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Lambiris, Elias J
;; REGISTRATION NUMBER: 33,728
;; REFERENCE/DOCKET NUMBER: 4944.200-US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-867-0123
;; TELEFAX: 212-878-9655
;; INFORMATION FOR SEQ ID NO: 12:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 524 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: No. 5958727e
US-08-928-692-12
Query Match 20.7%; Score 575.5; DB 2; Length 524;
Best Local Similarity 33.1%; Pred. No. 8.8e-47;
Matches 136; Conservative 94; Mismatches 147; Indels 34; Gaps 11;
QY 132 NLSISEPSAFGSLNVGMVGAISQVAEYIGRKSIMIAAIPNIIGWLAIAPAK---D 188
DB 98 SLSVSSFA-----VGGMTASFFGGWLGDTLGRKAKLVANILSLVGALLMGFSKLGPS 150
QY 189 ASFLVMGRLLGEGFGVGIISYTPVYIAEISPPQNRGALGSVNOLSVTFGIF-----LAY 242
DB 151 HILLIAGRSISGLYCGILSGLVPMYIGEIAFTALRGALGTFHQLAIVTGLISQIIGLEF 210
QY 243 LLGMPIWRLLAVIGALPCTNMLIPGLFFIPESPRWL-AKXNLTDCETSQVLRGPETDI 301
DB 211 ILGNVYLWHILLGSGVRAILQSLLLFFCPESPRLYIKLDEEVKAKQSLKRLRGYD-DV 269
QY 302 TTEVNDIK--RAVASSSKRTTISFOELNOKKYRTEPLLGLLGLVLQNLGSGINGVLV 359
DB 270 TKDINEMKREERESSEKQVSI-IQLFTNSSYQPIVLVALLMHLVAQOFSINGIFYSTS 328
QY 360 IFKAAGVNSDLATCSLGAIQVLAQTGT-TWLDRAGRRIILLIISTGTMCLLAVSVVFF 419
DB 329 IFQTAGISKPVYATIGVAVNMVFTAVSVLVEKAGRSFLIGMSGMFVCAIFMSVGLV 388
QY 420 LKDNISQDSNYYLTMTSLVGVSVFVITFSFGNGAIPWLMWSEILPVSIKSLGGSATL 479
DB 389 L-----LNKFSWMSYVSMIAIFLVSPFEIGPGPIPMFWVAEFPSSQGRPALAIAAF 441
QY 480 ANWLTSPAFTWTTNMLTWSVGTFSLSYM---VVSFTIVFVVLWVPETKG 527
DB 442 SNWTCNFIVALCFQVIADFC--GPYVFFLFAVLALFTL-FTFKVPEPKG 489

RESULT 3
US-09-339-972-12
; Sequence 12, Application US/09339972
; Patent No. 6323002
; GENERAL INFORMATION:
; APPLICANT: Brody, Howard
; APPLICANT: Yaver, Deborah S.
; APPLICANT: Lamsa, Michael
; APPLICANT: Hansen, Kim
; TITLE OF INVENTION: Methods for Modifying the Production of
; TITLE OF INVENTION: a Polypeptide
; NUMBER OF SEQUENCES: 80
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6323002o No. 6323002disk of No. 6323002th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:


```

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/339,972
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/928,692
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4944.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 524 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6323002e
; US-09-339-972-12

Query Match 20.7%; Score 575.5; DB 3; Length 524;
Best Local Similarity 33.1%; Pred. No. 8.8e-47;
Matches 136; Conservative 94; Mismatches 147; Indels 34; Gaps 11;

QY 132 NLGTSFSAFGSLSNVGMGVAIASQMAEYIGRKGSLMTAAIPNIGWLAISPAK---D 188
Db 98 SLVSSEFA-----VGMTPASFFGWGLGDTLGRKIKAMLVANILSLVGLLGMGFSKLGPS 150

QY 189 ASFLYMGRLLEGFGVGLISTVPYIAEISPNQMRGALGSVNQLSVTFGIF-----L 242
Db 151 HLIIAGRSISGLYGLISGLVPNYIGEIAPTALRGALGTFHQLAIVTGLISQIIGLEF 210

QY 243 LIGMFIPIWLLAVIGALPCTMLIPGLFFIPESPRWL-AKONLTEDCETSLQVLRGFETDI 301
Db 211 ILGNYDLWHLLGSLGVRAILQSLLLPFCPSRYLYIKLDEEVKAKQSLKRLGYD-DV 269

QY 302 TTEVNDIK-RAVASSKRTTISQELNOKKYRTPLLLGILGLVLQNLGSLNGVLFVASS 359
Db 270 TKDINEMKEREASSEQKSVI-IQLFTNSSYRQPIILVALMLHVAQQFSGINGIFYSTS 328

QY 360 IFKAGVTNSDLATCSLGAIOVLATGVTTLLDRAGRRIILITSTSGMTCLLAVSVVFF 419
Db 329 IFQTAGISKPVYATIGVAGVNVFTAVSVFLVEKAGRRSLFLIGMSGNFVCAIFMSVGLV 388

QY 420 LKDNISQDSNYIYLTWISLVGIVSFVITFSFGMGAIPWLMMSILPVSIKSLGGSIATL 479
Db 389 L-----LNKFSWMSYVSMIAIFLVSFPEIGPCIPFWFWVABFFSQGPRPAALAI 441

QY 480 ANWLTSAITMTNMLTWSVGGFLSYM-----VVSASTIVFVVLWVPETKG 527
Db 442 SNWTCNFVALCFQYIADFC--GPYVFLFAGVLLAFTL-FTFFKVPETKG 489

RESULT 4
US-09-031-392-5
; Sequence 5, Application US/09031392
; Patent No. 5942398
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street

; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-031-392-5

Query Match 20.6%; Score 572.5; DB 2; Length 494;
Best Local Similarity 32.2%; Pred. No. 1.6e-46;
Matches 147; Conservative 86; Mismatches 173; Indels 51; Gaps 14;

QY 105 IVALGPIQGFSTG-FSSPTQDAMVRD-LNLSISE-----FSAPGSL-----NVG 148
Db 16 IATIGSFQGVNTGVINAP--EAIKDFLNTLBERSETPPSSVLLTSLWSLVAIFS 73

QY 149 GMVGAIASQMAEYIGRKGSLMTAAIPNIGWLAISPAKDA---SFLYMGRLLEGFGVGI 205
Db 74 GMIGFSVGLFVNRGRRNSMLVNLALAGCLMGFCIAESVEMVLLGLILGLFCGL 133

QY 206 ISYTPVPIVIAEISPNQMRGALGSVNQLSVTFGIFLAILLGMPI-----PWLLAVIGAL 259
Db 134 CTGFVPMYIGEISPTALRGALGFTLNQLGIVIGILVAQIFGLKVLGTEDLWPLLLGFTIL 193

QY 260 PCTMLIPGLFFIPESPRMLAKNLTED--CETSLQVLRGFETDITTEVNDIKRAVASSSK 317
Db 194 PAIIQCAALPFCPSPRFLL-INRKEEKAKILORLWGTEDVAQDQEMKDESMRMSQ 251

QY 318 RTTISQDEL-NOKKYRTPLLLGILGLVLQNLGSLNGVLFYASSIFKAGVTNSDLATCSL 376
Db 252 EKQVTVLELFRAPNVRQPIIISIMQLSQLSGINAVFYSTGIFKDAQVQEPVATIGA 311

QY 377 GAIQVLAQVTTWLLDRAGRRIILITSTSGMTCLLAVSVVFLKDNISQDSNSYIILTM 436
Db 312 GVVNTIFTVVSFLVERAGRRTLHILGLGMAFCSILMTISLLLNKDN-----YSWMSF 364

QY 437 ISLVGIVSFVITFSFGMGAIPWLMMSILPVSIKSLGGSIATLANWLTSAITMTNML 496
Db 365 ICIGALLVFAVEFGPGPIFWFVIAELFGQGPRAAMAVACGSNWTNSFLVG-----L 418

QY 497 TWSVGGTFLSYMVVSASTIVFVVLW-----VPETKG 527
Db 419 LFPSATFYLGAVFVIVFTVFLVFWVFTFFKVPETRG 455

RESULT 5
US-09-299-549-5
; Sequence 5, Application US/09299549
; Patent No. 6136547
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
```

```

; TITLE OF INVENTION: GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,549
; FILING DATE: 26-APR-1999
; PRIOR APPLICATION NUMBER:
; FILING DATE: 09/031,392
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-299-549-5

```

```

Query Match      20.6%; Score 572.5; DB 3; Length 494;
Best Local Similarity 32.2%; Pred. No. 1.6e-46;
Matches 147; Conservative 86; Mismatches 173; Indels 51; Gaps 14;

QY 105 IVALGPTQFGTSG-FSSPTQDAMVRD-LNLISSE-----PSAFGSL-----NVG 148
DB 16 IATIGSFQFGYNTGVINAP--EAIKDFLNYTLERSETPPSSVLLTSLWSLSVAIFSVG 73

QY 149 GMVCAIASGQMAEYIGRKSIMIAIPNIIGWLAIISPAKDA---SFLYMGRLLEGFGVGI 205
DB 74 GMIGSFVGLFVNRFGRRNSMLIWNLLAIAGGCLMGFCKIAESVEMILGLRIIGLFCGL 133

QY 206 ISYTVPVYIAEISPNQNRGALGSVNQLSVTFGIFLAYLLGMFI-----PWRLAVIGAL 259
DB 134 CTGFVPMYIGEISPTALRGAFGLNQLGIVIGLVIAQIFGLKVILGTEDLWPLLLGFTIL 193

QY 260 PCTMLIPGLFFIPESPRWLAKMNLTD--CETSLQVLRGFETDITTEVNDIKRAVASSSK 317
DB 194 PAIIQCAALPFCSPESPRFL--INRKEEKAKELQLRWGTE-DVAQDIQEMKDESVMRSQ 251

QY 318 RTTISFOEL-NQKYRTPLLGLGLVQLNSGVLVYASSIFKAAQVYNSDLATCSL 376
DB 252 EKQVTVLELFRAPNYRQPIIISIMLSQOLSGINAVFYSTGIFKDGAVQEPVYATIGA 311

QY 377 GAQVQLATGVTTWLLDRAGRILLIISTGMITLCLLAVSVWFFLKNIQSODNSVYILTM 436
DB 312 GVNTTFTVVSVELVERAGERTLHLGLGMAFCSILMTISLLKDN-----YSWMSF 364

QY 437 ISLVGVISVFTPSFGKGAIPWLMSEIIPVSIKSLGSGIATLANWLTSFAITMTNML 496
DB 365 ICIGAILVFAEPIGPGIPWFIIVAELEFGQGPRAAMAVAGCSNWTSTFLVG-----L 418

QY 497 TWSVGGTFLSMVVSFAITVIVVLW-----VPETKG 527
DB 419 LFPSATFYLGAVFYVFTVFLVIFWVFTFKVPEIRG 455

```

```

RESULT 6
US-09-610-417-5
; Sequence 5, Application US/09610417
; Patent No. 6346374
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/610,417
; FILING DATE: 05-Jul-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/299,549
; FILING DATE: <UNKNOWN>
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-610-417-5

```

```

Query Match      20.6%; Score 572.5; DB 3; Length 494;
Best Local Similarity 32.2%; Pred. No. 1.6e-46;
Matches 147; Conservative 86; Mismatches 173; Indels 51; Gaps 14;

QY 105 IVALGPTQFGTSG-FSSPTQDAMVRD-LNLISSE-----PSAFGSL-----NVG 148
DB 16 IATIGSFQFGYNTGVINAP--EAIKDFLNYTLERSETPPSSVLLTSLWSLSVAIFSVG 73

QY 149 GMVCAIASGQMAEYIGRKSIMIAIPNIIGWLAIISPAKDA---SFLYMGRLLEGFGVGI 205
DB 74 GMIGSFVGLFVNRFGRRNSMLIWNLLAIAGGCLMGFCKIAESVEMILGLRIIGLFCGL 133

QY 206 ISYTVPVYIAEISPNQNRGALGSVNQLSVTFGIFLAYLLGMFI-----PWRLAVIGAL 259
DB 134 CTGFVPMYIGEISPTALRGAFGLNQLGIVIGLVIAQIFGLKVILGTEDLWPLLLGFTIL 193

QY 260 PCTMLIPGLFFIPESPRWLAKMNLTD--CETSLQVLRGFETDITTEVNDIKRAVASSSK 317
DB 194 PAIIQCAALPFCSPESPRFL--INRKEEKAKELQLRWGTE-DVAQDIQEMKDESVMRSQ 251

QY 318 RTTISFOEL-NQKYRTPLLGLGLVQLNSGVLVYASSIFKAAQVYNSDLATCSL 376
DB 252 EKQVTVLELFRAPNYRQPIIISIMLSQOLSGINAVFYSTGIFKDGAVQEPVYATIGA 311

QY 377 GAQVQLATGVTTWLLDRAGRILLIISTGMITLCLLAVSVWFFLKNIQSODNSVYILTM 436
DB 312 GVNTTFTVVSVELVERAGERTLHLGLGMAFCSILMTISLLKDN-----YSWMSF 364

```

QY 437 ISLVGVSVFVITFSGMGAIPWLMSEILPVSIKSLGGSATIANWLTSAFTMTNML 496
Db 365 ICIGAILVFAFPEIGPGIPWFIIVAEFLPGQGPRAAMAVAGCSNNTSFLVG-----L 418
QY 497 TWSVGGTFLSYMVVSFAFTIVFVLM-----VPETKG 527
Db 419 LPFSAFYLGAYVFIIVTFVFLVIFVFWTFPKVPETKG 455

RESULT 7
US-09-031-392-10
; Sequence 10, Application US/09031392
; Patent No. 5942398
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 493 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-031-392-10

Query Match 20.4%; Score 568; DB 2; Length 493;
Best Local Similarity 30.9%; Pred. No. 4.2e-46;
Matches 142; Conservative 87; Mismatches 182; Indels 48; Gaps 9;

QY 101 LCTLVAGLPQFGTSGF-----SSPTQDAMVRDLNLSIFE 140
Db 16 LAVLIAALGSFOGYNGLVGNAPQKIEAFYETWLGKXGXPSPVPTLLNLSVSIFA- 74

QY 141 FGSLSNVGMVGAISQMAEYIGRKGSLMIAIPNIIGWLAIISFAKDA---SFLYMGRL 197
Db 75 -----VGMIGSFLVGVXIGNLGRKXKXAMLVNVLAIAGLLMGLAKXAXSPFEMILGRF 128

QY 198 LEGFGVGIISYTVPIVYIAEISQPNMRGALGSVNQLSVTFGFIPLAYLLGM-----FIPWR 251
Db 129 IIGLYCGLSSGVPMYVGEISPTALRGALGTNLQGLIGVIGILIAQVLGLDLSLGNESLWP 188

QY 252 LLAVIGALPCTMLIPGLFFPIEPESPRWLAKNLTED--CETSLQVLKGFETDITEYNDIK 309
Db 189 LLLGLTGPALLQLLLPFCPEPRYLL-INKEEAPAKALQRLG-TADVSGQVAEYK 246

QY 310 RAVASSKRTTISQEL-NQKYRTPLLLGILLVLQNLGSLNGVLFYASSIFPAAGVTN 368
Db 247 DESRXMXSERKXSVLELFRGRXYRQPVIIAIVLQLSQQLSGINAVFYSTSIPEKAGVGQ 306

QY 369 SDLATCSLGAIOVLATGVTWLLDRAGRILLIISTGWTICLLAVSVVFLKDNISQDS 428
Db 307 PVYATIGAGVNTVFTVSVFVVERAGRTLHLGLGGMAGCAVMTTIALALLDQVPW-- 364
QY 429 NSYVILTMISLVGVSVFVITFSGMGAIPWLMSEILPVSIKSLGGSATIANWLTSAF 488
Db 365 -----MSYVSIVAIFGVFAFEVGPPIPWFIIVAEFLSQGPRAAIAVAGFSNNTSFEIV 419

QY 489 TMTNMLTWSVGGTFLSYMVVSFAFTIVFVLMVPETKG 527
Db 420 GLLFOYIAELLGPYVFIIVFAVLLLLFFITFLKVPETKG 458

RESULT 8
US-09-299-549-10
; Sequence 10, Application US/09299549
; Patent No. 6136547
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,549
; FILING DATE: 26-APR-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/031,392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 493 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-299-549-10

Query Match 20.4%; Score 568; DB 3; Length 493;
Best Local Similarity 30.9%; Pred. No. 4.2e-46;
Matches 142; Conservative 87; Mismatches 182; Indels 48; Gaps 9;

QY 101 LCTLVAGLPQFGTSGF-----SSPTQDAMVRDLNLSIFE 140
Db 16 LAVLIAALGSFOGYNGLVGNAPQKIEAFYETWLGKXGXPSPVPTLLNLSVSIFA- 74

QY 141 FGSLSNVGMVGAISQMAEYIGRKGSLMIAIPNIIGWLAIISFAKDA---SFLYMGRL 197
Db 75 -----VGMIGSFLVGVXIGNLGRKXKXAMLVNVLAIAGLLMGLAKXAXSPFEMILGRF 128

QY 198 LEGFGVGIISYTVPIVYIAEISQPNMRGALGSVNQLSVTFGFIPLAYLLGM-----FIPWR 251
Db 129 IIGLYCGLSSGVPMYVGEISPTALRGALGTNLQGLIGVIGILIAQVLGLDLSLGNESLWP 188

QY 314 SSSKRT-----TISFOELNQKRYTPLLIGLVLQNLGSLGVLFYASSIFKAAG 365
DB 286 SEKRVLKOLILSPPTAMRH-----LIAGIGHFFQSSGIDAVVLYSPLVFXSAG 337
QY 366 VTNSDL---ATCSLGAIOVLATQVTTWLLDRAGRILLIISTSCMTLCLLAVSVVFFLKD 422
DB 338 ITGDSRLRGTTAVAGATNTVFILVATFLDLDRRRLPVLITSTGMLVSLVGLATGLTVIS 397
QY 423 NISODNSYILTMISLNGIVSFVTFSGMGAIPWLMSEILPVSIGSLGGSIAFLANN 482
DB 398 RHDEKITWAI--VLICFCIMAVAFSGLGLPTWVYSSEIFPLHVRALGCSLGVAVNR 455
QY 483 LTSFAITMT-TNMLTWSVGGTFLSYMVVSFAFTIVFVWLWVPETKG 527
DB 456 LTGCVISMTFISLKAHTIGGAFFLFAGIASFAWVFFAYLPBTRG 501

RESULT 11
US-09-031-392-6
; Sequence 6, Application US/09031392
; Patent No. 5942398
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072001
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 509 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-031-392-6

Query Match 19.8%; Score 550; DB 2; Length 509;
Best Local Similarity 30.4%; Pred. No. 2.5e-44;
Matches 142; Conservative 90; Mismatches 177; Indels 58; Gaps 9;

QY 103 TLIVA-----LGPIQFGFTSGFSSPTQDAMVRDLN----- 132
DB 23 TLVLAVSFVLSLQFGYGNIGVINAPQVIEQSNETWLGQEGPSSIPPGLTTLWA 82
QY 133 LSISEFAGFSLSNVGMWGIAGQMAVEYIGRKGSLMIAIPNIGWLAIKAKAS-- 190
DB 83 LSAVIFS-----VGMISFLIGIISQWLGKRAMLVNVLAVLGGSLMGLANAASY 135
QY 191 -FLYMGRLLEGFGVGIISYTPVPIABISPNOMEGALGSVNOLSVTFGIFLAY-----L 243

DB 136 EMLILGRFLICAYSLTSLVPMVYGEIAPHLRGALGTNLQNLAIIVIGILLAQVLGLES 195
QY 244 LGMFIPWRLLAIVGALPCTMLIPGLFFIPESPRWL-AKONJTEBCESTSLQVLRGFE--TD 300
DB 196 LGTASLWPLLLGLVLPALLQVLVLLPFCPESPRYLYITQLEGPAKSLKELTGWADVSG 255
QY 301 ITTEVNDIKRAVASSKETTISFOELNQKRYTPLLIGLVLQNLGSLGVLFYASSI 360
DB 256 VLAEKDEKRL--ERERPLSLQLLQSGRHRQRELIIVLQLSQQLSGINAVFYSTSI 313
QY 361 FKAAGVTNSDLATCSLGAIOVLATQVTTWLLDRAGRILLIISTSCMTLCLLAVSVVFFL 420
DB 314 FETAGVGQPAYATIGAGVNTVFTLVSVLLVERAGRRTTHLLGLAGMCGCAILMTVALL 373
QY 421 KDNISQDSNSYIILTMISLNGIVSFVTFSGMGAIPWLMSEILPVSIGSLGGSIAFLA 480
DB 374 LERVE-----AMSYVSIIVAFGFAFFEGPPIPMFIVAEFLVSQGPRAAMAVAGFS 426
QY 481 NWLTSFAITMTNMLTWSVGGTFLSYMVVSFAFTIVFVWLWVPETKG 527
DB 427 NWTSNFIICMGFOYVAEMGPYVLLFAVLLGLGFFITFLVLPBTRG 473

RESULT 12
US-09-299-549-6
; Sequence 6, Application US/09299549
; Patent No. 6136547
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,549
; FILING DATE: 26-APR-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/031,392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 509 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-299-549-6

Query Match 19.8%; Score 550; DB 3; Length 509;
Best Local Similarity 30.4%; Pred. No. 2.5e-44;
Matches 142; Conservative 90; Mismatches 177; Indels 58; Gaps 9;

QY 103 TLIVA-----LGPIQFGFTSGFSSPTQDAMVRDLN----- 132

Db 23 TLVAVFSAVLGSLQFGYINIGVINAPQKVIQSYNETWLGROGPEGSSIPPCTLTTLWA 82
 QY 133 LSISEFAFGSLNVGVMGAIASQMAEYIGRKSLMIAAIPNIIGWLAIISFAKAS-- 190
 Db 83 LSAVAFS-----VGMISFLIGIISQWLGRKRAMLVNNVLAIVLGSLMGLANAAS 135
 QY 191 -FLVMGRLLGFGYGIISYTPVVIARISPNQMRGALGSVNQLSVTFGIFLAY-----L 243
 Db 136 EMLILGRFLIGAYSLGSLVPMVYGEIATHLRGALGTNLQALVIGILIAQVLGLES 195
 QY 244 LGMFIPRLLAVIGALPCTMLIPGLFPIPSRWL-AKMLTDEDCETSLQVLRGE--TD 300
 Db 196 LGTASLMPLLGLTVPALLQVLLPFCPSRYLYIQNLEGPARKSLKRLTGWADVSG 255
 QY 301 ITTEVNDIKRAVASSKRTTISFOELNOKKYRTPLLLIGLLVQLNLSINGVLFYASSI 360
 Db 256 VLAEKDEKRL--ERERPLSLQLGSRTHRQPLIAVAVLQSLGINAVFYYSI 313
 QY 361 FKAAGVTNSDLATCSLGAIOVLATGVTTWLLDRAGRILLIISTSGMTCLLAVSVVFFL 420
 Db 314 FETAGVGPAYATIGAGVWNTVTLVSVLLVERAGRTLHLGLAGMCGCAILMTVALL 373
 QY 421 KDNISQDSNSYILTMISLGVISFVITFSFGMAIPWLMMSSEILPVSIKSLGSIATLA 480
 Db 374 LERVP-----AMSVYSIVAIFGVAFPIGPIWFIIVAEFSGQPRPAMAVAGFS 426
 QY 481 NWLTSFAITMTNMLTWSVGGTFLSYMVVSFAFTIVFVVLWVPETKG 527
 Db 427 NWTNFIIGMGFQVVAEAMGPYVFLFAVLLGGFFITFLRVPETRG 473

RESULT 13

US-09-610-417-6
 ; Sequence 6, Application US/09610417
 ; Patent No. 6346374
 ; GENERAL INFORMATION:
 ; APPLICANT: Tarraglia, Louis A.
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
 ; GLUTEX AND USES THEREOF
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows95
 ; SOFTWARE: FASTSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/610,417
 ; FILING DATE: 05-Jul-2000
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/299,549
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Meiklejohn, Ph.D., Anita L.
 ; REGISTRATION NUMBER: 35,283
 ; REFERENCE/DOCKET NUMBER: 07334/072002
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617/542-5070
 ; TELEFAX: 617/542-8906
 ; TELEX: 200154
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 509 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein

US-09-610-417-6
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
 Query Match 19.8%; Score 550; DB 3; Length 509;
 Best Local Similarity 30.4%; Pred. No. 2.5e-44;
 Matches 142; Conservative 90; Mismatches 177; Indels 58; Gaps 9;
 QY 103 TLIVA-----LGTQFGFTSGFSSPTODAMVRDLN----- 132
 Db 23 TLVAVFSAVLGSLQFGYINIGVINAPQKVIQSYNETWLGROGPEGSSIPPCTLTTLWA 82
 QY 133 LSISEFAFGSLNVGVMGAIASQMAEYIGRKSLMIAAIPNIIGWLAIISFAKAS-- 190
 Db 83 LSAVAFS-----VGMISFLIGIISQWLGRKRAMLVNNVLAIVLGSLMGLANAAS 135
 QY 191 -FLVMGRLLGFGYGIISYTPVVIARISPNQMRGALGSVNQLSVTFGIFLAY-----L 243
 Db 136 EMLILGRFLIGAYSLGSLVPMVYGEIATHLRGALGTNLQALVIGILIAQVLGLES 195
 QY 244 LGMFIPRLLAVIGALPCTMLIPGLFPIPSRWL-AKMLTDEDCETSLQVLRGE--TD 300
 Db 196 LGTASLMPLLGLTVPALLQVLLPFCPSRYLYIQNLEGPARKSLKRLTGWADVSG 255
 QY 301 ITTEVNDIKRAVASSKRTTISFOELNOKKYRTPLLLIGLLVQLNLSINGVLFYASSI 360
 Db 256 VLAEKDEKRL--ERERPLSLQLGSRTHRQPLIAVAVLQSLGINAVFYYSI 313
 QY 361 FKAAGVTNSDLATCSLGAIOVLATGVTTWLLDRAGRILLIISTSGMTCLLAVSVVFFL 420
 Db 314 FETAGVGPAYATIGAGVWNTVTLVSVLLVERAGRTLHLGLAGMCGCAILMTVALL 373
 QY 421 KDNISQDSNSYILTMISLGVISFVITFSFGMAIPWLMMSSEILPVSIKSLGSIATLA 480
 Db 374 LERVP-----AMSVYSIVAIFGVAFPIGPIWFIIVAEFSGQPRPAMAVAGFS 426
 QY 481 NWLTSFAITMTNMLTWSVGGTFLSYMVVSFAFTIVFVVLWVPETKG 527
 Db 427 NWTNFIIGMGFQVVAEAMGPYVFLFAVLLGGFFITFLRVPETRG 473

RESULT 14

US-09-291-922-22
 ; Sequence 22, Application US/09291922
 ; Patent No. 6383776
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Hitz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB-1163
 ; CURRENT APPLICATION NUMBER: US/09/291,922
 ; CURRENT FILING DATE: 1999-04-14
 ; EARLIER APPLICATION NUMBER: 60/083,044
 ; EARLIER FILING DATE: April 24, 1998
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 22
 ; LENGTH: 510
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; NAME/KEY: UNSURE
 ; LOCATION: (102)
 ; US-09-291-922-22
 Query Match 19.8%; Score 550; DB 3; Length 510;
 Best Local Similarity 29.6%; Pred. No. 2.5e-44;
 Matches 141; Conservative 97; Mismatches 201; Indels 38; Gaps 8;
 QY 99 AFLCTLIVAGLPQFGFTSGFSSPTQDAMVRDLNLSISFSAFGSLNVGVMGAIASGQ 158
 Db 21 AFACAILASWTSILLGYDVGMSGASLYIKKDNISDKVEVLGMLNLSLIGSFAAGR 80

THIS PAGE BLANK (uspio)

OM protein - protein search, using sw model

Run on: October 13, 2004, 11:38:38 ; Search time 209.265 Seconds
(without alignments)
852.715 Million cell updates/sec

Title: US-10-051-909-36
 Perfect score: 2779
 Sequence: 1 PSSSSSFRPAGKKKKKKNQG.....TIFVSLSIQRLQWLPECLS 553

Scoring table: BLOSUM62

Searched: 1356558 seqs, 322682953 residues

Total number of hits satisfying chosen parameters: 1356558

Minimum DB seq length: 0

```

Maximum DB seq range: 200000000

```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

```
Database :
Firmware: Applications
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pcp
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pcp
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pcp
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pcp
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pcp
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pcp
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pcp
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pcp
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pcp
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pcp
18: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp
19: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pcp
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB	ID	Description
1	2777	99.9	553	13	US-10-051-909-36	Sequence 36, Appl
2	2119.5	75.9	523	15	US-10-435-114-63431	Sequence 63431, A
3	2108.5	76.9	501	16	US-10-437-963-163544	Sequence 163544, A
4	1801.5	64.8	501	13	US-10-051-909-38	Sequence 38, Appl
5	1796.5	64.6	509	15	US-10-435-114-59933	Sequence 59933, A
6	1718	51.8	484	15	US-10-424-599-275264	Sequence 275264, A
7	1581	51.8	489	15	US-10-435-114-63430	Sequence 63430, A
8	1378	49.6	414	16	US-10-437-963-165948	Sequence 165948, A
9	1219	43.9	318	15	US-10-435-114-65425	Sequence 65425, A
10	1199	43.1	361	15	US-10-434-599-275268	Sequence 275268, A
11	1185	42.6	338	15	US-10-434-599-273908	Sequence 273908, A
12	1032	37.1	219	15	US-10-435-114-49786	Sequence 49786, A
13	909.5	32.7	502	16	US-10-437-963-166774	Sequence 166774, A
14	891	32.1	481	14	US-10-310-154-725	Sequence 725, App

15	768	27.6	216	15	US-10-424-599-275267	Sequence 275267, A
16	759	27.3	54	16	US-10-437-963-106607	Sequence 106607, A
17	758.5	27.3	263	15	US-10-424-599-275265	Sequence 275265, A
18	717.5	25.8	421	15	US-10-424-599-347207	Sequence 247207, A
19	658.5	23.7	477	14	US-10-168-661-1	Sequence 1, Appli
20	675.5	23.7	477	9	US-09-886-954-1	Sequence 1, Appli
21	635.5	22.5	228	15	US-10-425-114-47622	Sequence 17622, A
22	623	22.4	432	15	US-10-023-020-186	Sequence 186, App
23	618	22.2	507	14	US-10-157-031-359	Sequence 359, App
24	618	22.2	507	16	US-10-755-889-110	Sequence 110, App
25	612.5	22.0	487	9	US-09-795-693-27	Sequence 27, Appl
26	612.5	22.0	487	13	US-10-095-139-14	Sequence 14, Appl
27	612.5	22.0	487	14	US-10-152-239-27	Sequence 27, Appl
28	612.5	22.0	487	14	US-10-199-485-27	Sequence 27, Appl
29	612.5	22.0	487	14	US-10-391-399-45	Sequence 45, Appl
30	609	21.9	486	9	US-09-860-232A-7	Sequence 7, Appli
31	607	21.8	488	10	US-09-794-832-11	Sequence 11, Appl
32	607	21.8	488	13	US-10-094-059-4	Sequence 4, Appli
33	607	21.8	488	14	US-10-170-538-5	Sequence 5, Appli
34	607	21.8	488	14	US-10-162-012-46	Sequence 46, Appl
35	607	21.8	488	14	US-10-063-960B-4	Sequence 4, Appli
36	607	21.8	488	14	US-10-144-634-4	Sequence 46, Appl
37	607	21.8	488	14	US-10-163-102-46	Sequence 4, Appli
38	607	21.8	488	15	US-10-038-854-170	Sequence 170, App
39	605.5	21.8	480	9	US-09-734-569-172	Sequence 172, App
40	601.5	21.6	157	16	US-10-767-701-50013	Sequence 50013, A
41	600.5	21.6	191	15	US-10-423-114-61925	Sequence 61925, A
42	599	21.6	161	15	US-10-425-114-49791	Sequence 49791, A
43	597.5	21.5	389	16	US-10-767-701-44031	Sequence 44031, A
44	592.5	21.3	535	9	US-09-795-693-20	Sequence 20, Appl
45	583.5	21.3	535	14	US-10-156-233-20	Sequence 20, Appl

ALIGNMENTS

```

RESULT 1
US-10-051-909-36
; Sequence 36, Application US/10051909
; Publication No. US20020-99217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: B0163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 36
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (528)
; OTHER INFORMATION: Xaa = ANY AMINO ACID
US-10-051-909-36

```

```
Query Match          99.9%; Score 2777; DB 13; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.5e-232;
Matches 553. Conservative 0; Mismatches 0; Indels 0; Caps 0;
```

Qy 1 PSSSSFPAGKKKKKQGLRRAVPGRPASLRTRVMGGSNRGAGAGEEGSDHG 60

61 VI.RRPLINTGSWYRMSSROSSFAPGTSSMAVLRESHVSAFLCTLIVALGPIQFGFTSGFS 120

Db 61 VLRPLNTGWSYMSRQSSFAFCTSMVLRSHVSAFLCTLIVAGLPQFGFTSGFS 120
Qy 121 SPTQDAMVRDLNLISBPSAFGSLSNVGMVGATASQMAEYIGRKSLMIAAIPNIGW 180
Db 121 SPTQDAMVRDLNLISBPSAFGSLSNVGMVGATASQMAEYIGRKSLMIAAIPNIGW 180
Qy 181 LAISFAKDASFLYGRLLLEGFGVGIISYTVVPVYIAEISPNQMRGALGSVNQLSVTFGIFL 240
Db 181 LAISFAKDASFLYGRLLLEGFGVGIISYTVVPVYIAEISPNQMRGALGSVNQLSVTFGIFL 240
Qy 241 AYLLGMFIPWLLAVIGALPCTMLIPGLFPTBPSPRLAKNLTDCETSLOVLRGFETD 300
Db 241 AYLLGMFIPWLLAVIGALPCTMLIPGLFPTBPSPRLAKNLTDCETSLOVLRGFETD 300
Qy 301 ITTEVNDIKRAVASSSKRTTISFQELNOKKYRTPLLLGIGLLVLQNLGSLNGVLFYASSI 360
Db 301 ITTEVNDIKRAVASSSKRTTISFQELNOKKYRTPLLLGIGLLVLQNLGSLNGVLFYASSI 360
Qy 361 PKRAGVNSDLATCSLGAIOVLATGVTTWLLDRGRILLIISTSGMTCLLAVSVVFFL 420
Db 361 PKRAGVNSDLATCSLGAIOVLATGVTTWLLDRGRILLIISTSGMTCLLAVSVVFFL 420
Qy 421 KDNISQDSNSYIITMISLGVISFVITTFSGMGAIPWLMMSILPVSIKSLGSIATLA 480
Db 421 KDNISQDSNSYIITMISLGVISFVITTFSGMGAIPWLMMSILPVSIKSLGSIATLA 480
Qy 481 NWLTSFAITMTNMLTWSVGGTFLSNVVSATFIVFVVLWVPETKXNSRGDTIFVSLS 540
Db 481 NWLTSFAITMTNMLTWSVGGTFLSNVVSATFIVFVVLWVPETKXNSRGDTIFVSLS 540
Qy 541 IOROLOWLPECLS 553
Db 541 IOROLOWLPECLS 553

RESULT 2

US-10-425-114-63431
; Sequence 63431, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 63431
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4371-017-Fl1_FLI.pep
US-10-425-114-63431

Query Match 76.3%; Score 2119.5; DB 15; Length 523;
Best Local Similarity 81.7%; Pred. No. 2.1e-175;
Matches 419; Conservative 41; Mismatches 36; Indels 17; Gaps 5;

Qy 40 GGGSNR-GGAGAGEESGDHD--GVLRRLPLL-NTGSWYMSRQSSSF-APGTSSMAVLRE 94
Db 20 GGAMNRRGGVGVGDESGSDYESGGNRPKLLMHTGWSYMRGSRQSLTACAGTSSMAVLRE 79
Qy 95 SHVSAFLCTLIVAGLPQFGFTSGFSPTQDAMVRDLNLISBPSAFGSLSNVGMVGAI 154
Db 80 SHVSAFLCTLIVAGLPQFGFTSGFSPTQDAMVRDLNLISBPSAFGSLSNVGMVGAI 139

Qy 155 ASQMAEYIGRKSGSLMIAAIPNIGWLAISFAKDASFLYGRLLLEGFGVGIISYTVVPYI 214
Db 140 ASQMAEYIGRKSGSLMIAAIPNIGWLAISFAKDASFLYGRLLLEGFGVGIISYTVVPYI 199
Qy 215 AEISPNQMRGALGSVNQLSVTFGIFLAVLGMFIPWLLAVIGALPCTMLIPGLFFIPES 274
Db 200 AEISPNQMRGALGSVNQLSVTFGIFLAVLGMFIPWLLAVIGALPCTMLIPGLFFIPES 259
Qy 275 PRLAKNLTDCETSLOVLRGFETDITTEVNDIKRAVASSSKRTTISFQELNOKKYRTP 334
Db 260 PRLAKNMTDDFETSLQVLRGFETDISAEVNDIKRAVASANKRTTIRFQELNOKKYRTP 319
Qy 335 LLLGIGLLVLQNLGSLNGVLFYASSIFKAAGVNNSDLATCSLGAIOVLATGVTTWLLDRA 394
Db 320 LTLGIGLLVLQNLGSLNGVLFYASSIFKAAGVNNSDLATCSLGAIOVLATGVTTWLLDRA 379
Qy 395 GRILLIISTSGMTCLLAVSVVFFLKDNIQDSNSYIITMISLGVISFVITTFSGMG 454
Db 380 GRILLIISTSGMTCLLAVSVVFFLKDNIQDSNSYIITMISLGVISFVITTFSGMG 439
Qy 455 AIPWLMMSILPVSIKSLGSIATLANWLTSTFAITMTNMLTWSVGGTFLSNVVSATF 514
Db 440 AIPWLMMSILPVSIKSLGSIATLANWLTSTFAITMTNMLTWSVGGTFLSNVVSATF 499
Qy 515 IVFVVLWVPETKXNSRGDTIFVSLSIQROLOW 547
Db 500 LVFVVLWVPETKXRT-----LEEIQW 520

RESULT 3

US-10-437-963-163544
; Sequence 163544, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 163544
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_62530C.1.pep
US-10-437-963-163544

Query Match 75.9%; Score 2108.5; DB 16; Length 501;
Best Local Similarity 81.9%; Pred. No. 1.8e-174;
Matches 417; Conservative 41; Mismatches 34; Indels 17; Gaps 5;

Qy 44 NR-CGAGAGEESGDHD--GVLRRLPLL-NTGSWYMSRQSSSF-APGTSSMAVLRESHVS 98
Db 2 NRGGGVGVGDESGSDYESGGNRPKLLMHTGWSYMRGSRQSLTACAGTSSMAVLRESHVS 61
Qy 99 AFLCTLIVAGLPQFGFTSGFSPTQDAMVRDLNLISBPSAFGSLSNVGMVGAIASQ 158
Db 62 AFLCTLIVAGLPQFGFTSGFSPTQDAMVRDLNLISBPSAFGSLSNVGMVGAIASQ 121
Qy 159 MAEYIGRKSGSLMIAAIPNIGWLAISFAKDASFLYGRLLLEGFGVGIISYTVVPYIAEIS 218
Db 122 MAEYIGRKSGSLMIAAIPNIGWLAISFAKDASFLYGRLLLEGFGVGIISYTVVPYIAEIS 181
Qy 219 PNQMRGALGSVNQLSVTFGIFLAVLGMFIPWLLAVIGALPCTMLIPGLFFIPESPRWL 278

Db 182 PNMKGALGVNSVTVGILLAYLGMFVWRLAVIGILPCTVLIPGLFFIPESPRWL 241
Qy 279 AKNLTEDCETSLQVLRGFTDITTEVNDIKRAVASSKRTTISFOELNOKKYRTPLLG 338
Db 242 AKNMMDDETSQVLRGFTDITTEVNDIKRAVASSKRTTISFOELNOKKYRTPLLG 301
Qy 339 IGLVLQNLGSGINGVLFYASSIFKAAGVTNSDLATCSLGAIVLATGVTWLLDRGRRI 398
Db 302 IGLVLQNLGSGINGVLFYASSIFKAAGVTNSDLATCSLGAIVLATGVTWLLDRGRRI 361
Qy 399 LLIISGWTCLLAVSVVFFKDNISODNSVYILTMISLVCIVSFVITESFGMGAI PW 458
Db 362 LLIISGWTCLLAVSVVFFKDNISODNSVYILTMISLVCIVSFVITESFGMGAI PW 421
Qy 459 LKMSILPVSIKSLGSGIATLANWLTSPAITMTTNLMTWSVGGTFLSYMVVSFAITVYV 518
Db 422 IIMSEILPVSIKSLGSGIATLANWLTSPAITMTTNLMTWSVGGTFLSYMVVSFAITVYV 481
Qy 519 VLVMPETKGNRSGDTTFVSLSTQRLQW 547
Db 482 ILWVPETKGR-----LEETQW 498

RESULT 4
US-10-051-909-38
; Sequence 38, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helencjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051.909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 38
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-051-909-38

Query Match 64.8%; Score 1801.5; DB 13; Length 501;
Best Local Similarity 72.2%; Pred. No. 8.8e-148;
Matches 351; Conservative 62; Mismatches 64; Indels 9; Gaps 2;
Qy 51 GEESGDHDG-----VLRRPLNTGWSYRMSRQSSFPAGT---SSNAVLRSHVSAFL 101
Db 5 GEESGDGGRASASDLRKPFHTGWSYRMSRQSSAGGGGWSRLGSSAYSURDSVSAVL 64
Qy 102 CTLI VALGPIQFGFTSGFSSPTQDAMVRDLNLSIFEAFSGSLNVGMVGAIASGMAE 161
Db 65 CTLI VALGPIQFGFTCGFSSPTQDAIISDLGLTSEFSLFGSLNVGMVGAIASGMAE 124
Qy 162 YIGKGSMLIAIPNIIGWLAIKFAKASFLYMGRLLEGFGVGIISITVPVYIAEISPN 221
Db 125 YIGKGSMLIAIPNIIGWLAIKFAKASFLYMGRLLEGFGVGIISITVPVYIAEISPN 184
Qy 222 MRGALGVNSVLTGFIPLAVLGMFIPWRLLAVIGALPCTWLPGLFFIPESPRWLAKM 281
Db 185 MRGALGVNSVLTGFIPLAVLGMFIPWRLLAVIGALPCTWLPGLFFIPESPRWLAKM 244
Qy 282 NLTEDETSQVLRGFTDITTEVNDIKRAVASSKRTTISFOELNOKKYRTPLLG 341
Db 245 GRWEDFESSQLVLRGFTDITTEVNDIKRAVASSKRTTIRFADIKOKKYSVPLMVIGL 304
Qy 342 LVLQNLGSGINGVLFYASSIFKAAGVTNSDLATCSLGAIVLATGVTWLLDRGRRI 401

Db 305 LVLQNLGSGINGVLFYASSIFKAAGVTNSDLATCSLGAIVLATGVTWLLDRGRRI 364
Qy 402 ISTGWTCLLAVSVVFFKDNISODNSVYILTMISLVCIVSFVITESFGMGAI PW 461
Db 365 ISTGWTCLLAVSVVFFKDNISODNSVYILTMISLVCIVSFVITESFGMGAI PW 424
Qy 462 SEILPVSIKSLGSGIATLANWLTSPAITMTTNLMTWSVGGTFLSYMVVSFAITVYV 521
Db 425 SEILPVSIKSLGSGIATLANWLTSPAITMTTNLMTWSVGGTFLSYMVVSFAITVYV 484
Qy 522 VPETKG 527
Db 485 VPETKG 490

RESULT 5
US-10-425-114-59933
; Sequence 59933, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59933
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3689-227-G11_FLI.pep
US-10-425-114-59933

Query Match 64.6%; Score 1796.5; DB 15; Length 509;
Best Local Similarity 70.4%; Pred. No. 2.4e-147;
Matches 352; Conservative 63; Mismatches 68; Indels 17; Gaps 4;
Qy 41 GGSNRGGAGAGEESG-----SDHDGVLRRPLNTGWSYR-----MSRQSSFPAGTSS 88
Db 3 GMSFRDQESGEGDGGRTSSASD---LRKPLNTGWSYRMPAGGVMGSRQSSLMERLGS 58
Qy 89 MAV-LRESHVSAFLCTLI VALGPIQFGFTSGFSSPTQDAMVRDLNLSIFEAFSGSLNV 147
Db 59 SAFSLRDVAISATLCTLI VALGPIQFGFTCGYSSPTQDAITADLGLSEFLFGSLNV 118
Qy 148 GGMVGAIASGMAEYIGKGSMLIAIPNIIGWLAIKFAKASFLYMGRLLEGFGVGIIS 207
Db 119 GAVVGAISGQLAEYIGKGSMLIAIPNIIGWLAIKFAKASFLYMGRLLEGFGVGIIS 178
Qy 208 YTPVYIAEISPNMRGALGVNSVLTGFIPLAVLGMFIPWRLLAVIGALPCTWLPGL 267
Db 179 YTPVYIAEISPNMRGALGVNSVLTGFIPLAVLGMFIPWRLLAVIGALPCTWLPGL 238
Qy 268 LFFIPESPRWLAKMNLTEDETSQVLRGFTDITTEVNDIKRAVASSKRTTISFOELN 327
Db 239 LFFIPESPRWLAKMNLTEDETSQVLRGFTDITTEVNDIKRAVASSKRTTIRFADIK 298
Qy 328 QKRYRTPLLGIGLLVLQNLGSGINGVLFYASSIFKAAGVTNSDLATCSLGAIVLATGVT 387
Db 299 QKRYRTPLLGIGLLVLQNLGSGINGVLFYASSIFKAAGVTNSDLATCSLGAIVLATGVT 358
Qy 388 TWLLDRAGRRIIISTGWTCLLAVSVVFFKDNISODNSVYILTMISLVCIVSFVI 447
Db 359 TWLLDRAGRRIIISTGWTCLLAVSVVFFKDNISODNSVYILTMISLVCIVSFVI 418

QY 448 TFSFGMGAIPMLMSEILPVSIKSGSIATLANLWTSFAITMTNMLTWSVGGTFLSY 507
Db 419 AFSLGLGAIPMLMSEILPVSIKSLAGSVATLANLWTAITMTASLANLWSSGTTAIY 478
QY 508 MVSAFTIVFVVLWVPETKG 527
Db 479 AVVSTMALIFVCLWVPETKG 498

RESULT 6

US-10-424-599-275264
; Sequence 275264, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 275264
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_90584C.1.pap
US-10-424-599-275264

Query Match 61.8%; Score 1718; DB 15; Length 484;
Best Local Similarity 69.5%; Pred. No. 1.5e-140;
Matches 331; Conservative 71; Mismatches 70; Indels 4; Gaps 2;
QY 52 EESGSDHGVLRRLTNGMYRMSRQSPFAPGCTSSMAVLRSHVSAFLCTLIIVAGPI 111
Db 2 EESGDARD--LQKFLHTGMYKMSQSSIM--GSTHVIRGANSVIFCVLIVAGPI 57
QY 112 QFGTSGFSSPTQDAMVRDLNLSISEPSAFGSLSNVGMVGAIASQMAFYIGRKGSLMI 171
Db 58 QFGTCGYSSPTQGAIVRDLNLSISEPSFPGSLSNVGMVGAIASQIAEYIGRKGSLMI 117
QY 172 AAIPIIIGWLAIAPKADASFLYGRLLGFGVGIISVTPVYIAEISPNMRGALGSVNO 231
Db 118 AAIPIIIGWLAIAPKADSSFLYGRLLGFGVGIISVTPVYIAEISPNMRGALGSVNO 177
QY 232 LSVTFGIFLAYLLGMFIPWRLAVIGALPCTMLIPGLFFIPESPRWLAKNLTDCETSL 291
Db 178 LSVTIGIMLAYLLGLFVNRVLAIGLPCTVLIPGLFFIPESPRWLAKNLTDCETSL 237
QY 292 QVLRGPTDITTEVNDIKRAVASSSKRTTISFQELNQKRYTPLLGLLVLQNLGSLN 351
Db 238 QVLRGPTDITTEVNDIKRAVASSSKRTTISFQELNQKRYTPLLGLLVLQNLGSLN 297
QY 352 GVLFYASSIFKAAGVNTSDIATCSLGAIOVLATGVTTWLLDRAGRILLIISTSGMTLC 411
Db 298 GILFYSTTIFANAGISSSEATVGLGAVQVIAIGISTVLVDKSGRLLIISTSGMTVSL 357
QY 412 LAVSVFPLKDNISQDSNSYIITMISLVGIVSFVITFSGMGAIPWLMSEILPVSIKSL 471
Db 358 LIVSIAPFYLEGVVSEDSHLFVSIIGVIVGLVAMVIGFSLGLPIPLWIMSEILPVN 417
QY 472 LGSIIATLANLWTSFAITMTNMLTWSVGGTFLSYMVVSAFTIVFVVLWVPETKG 527
Db 418 LAGSIATGNWLLISWGTMTANLLWSSGGTFTIVTVVAAFTIAPVWVPETKG 473

RESULT 7

US-10-425-114-63430
; Sequence 63430, Application US/10425114
; Publication No. US20040034888A1
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966

; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 63430
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4309-039-E8_FLI.pap
US-10-425-114-63430

Query Match 56.9%; Score 1581; DB 15; Length 489;
Best Local Similarity 82.8%; Pred. No. 1.3e-128;
Matches 308; Conservative 29; Mismatches 23; Indels 12; Gaps 1;
QY 176 NIIGWLAIAPKADASFLYGRLLGFGVGIISVTPVYIAEISPNMRGALGSVNO 235
Db 127 NIIGWLAIAPKADSSFLYGRLLGFGVGIISVTPVYIAEISPNMRGALGSVNO 186
QY 236 EGIPLAYLLGMFIPWRLAVIGALPCTMLIPGLFFIPESPRWLAKNLTDCETSL 295
Db 187 VGILLAYLLGMFIPWRLAVIGALPCTMLIPGLFFIPESPRWLAKNLTDCETSL 246
QY 296 GFETDITTEVNDIKRAVASSSKRTTISFQELNQKRYTPLLGLLVLQNLGSLN 355
Db 247 GFETDISAEVNDIKRAVASSSKRTTIRFQELNQKRYTPLLGLLVLQNLGSLN 306
QY 356 YASSIFKAAGVNTSDIATCSLGAIOVLATGVTTWLLDRAGRILLIISTSGMTLC 415
Db 307 YAGSIFKAAGVNTSDIATCSLGAIOVLATGVTTWLLDRAGRILLIISTSGMTLC 366
QY 416 VVFLKDNISQDSNSYIITMISLVGIVSFVITFSGMGAIPWLMSEILPVSIKSL 475
Db 367 VVFLKDNISQDSNSYIITMISLVGIVSFVITFSGMGAIPWLMSEILPVSIKSL 426
QY 476 IATLANWTSFAITMTNMLTWSVGGTFLSYMVVSAFTIVFVVLWVPETKG 535
Db 427 FATLANWTSFAITMTNMLTWSVGGTFLSYMVVSAFTIVFVVLWVPETKG 480
QY 536 FVLSLQRLQW 547
Db 481 -----LEEIQW 486

RESULT 8

US-10-437-963-165948
; Sequence 165948, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barabazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966

```
; SEQ ID NO 165948
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(414)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_64704C.1.pep
US-10-437-963-165948

Query Match
Best Local Similarity 49.6%; Score 1378; DB 16; Length 414;
Matches 289; Conservative 46; Mismatches 55; Indels 96; Gaps 6;

QY 51 GESSGSDHGDG-----VLREPLNTGSGWYMSRQSGSFAPGT---SSMAVLRSHVSAFL 101
Db 5 GESSGEDGRTASDLRPFLHTGSKWYKSSAGCGGGGSLGSSAYSLRDSVSAVL 64
QY 102 CTLIVAGLPQFGFTGSGFSDPTQDAMVRDNLNLSISFSAFGSLSNVGGVMGAIASQMAE 161
Db 65 CTLIVAGLPQFGFTGSGFSDPTQDAISDGLTLSEFSLFGLSLNVGMVGAIASQIAE 124
QY 162 YIGKSGSLMTAAIPNIIGLWLAISFAKDASFLYNGRLLEGFGVGIISYTPVPIAEISPN 221
Db 125 YIGKSGSLMTAAIPNIIGLWLAISFAKDS-----XVPVPIAEISAPQT 165
QY 222 MRGALGSVNQLSVTFGLFAYLGMFIPWRLAVIGALPCTMLIPGLFFIPESPRWLAKM 281
Db 166 MRGALGSVNQLSVTIGLLAYLGMFVFWPILSVGLPILGILPGLFFIPESPRWLAKM 225
QY 282 NLTFDCETSLQVLRGFTDITTEVNDIKRAVASSKRTTISFOELNQKVRTPLLLIGL 341
Db 226 GKMEDFESSQLVLRGFTDITTEVNDIKRAVASSKRTTISFOELNQKVRTPLLLIGL 258
QY 342 LVQLNLSGNGVLFYAGSIFKAAGVTNSDLATCSLGAIQVLTATGVTWLLDRAGRILLI 401
Db 259 LVQLNLSGNGVLFYAGSIFKAAGVTNSDLATCSLGAIQVLTATGVTWLLDRAGRILLI 297
QY 402 ISTGMDTCLLAVSVFLLKDNISQDSNSVYITMTSLVGIVSFVITFSGMGAIPLWLM 461
Db 298 ISTGMDTCLLAVSVFLLKDNISQDSNSVYITMTSLVGIVSFVITFSGMGAIPLWLM 339
QY 462 SEILPVSIKSGSIATLANWLTSPAITMTNLTMTWSVGTSLSYMVSAFTIVFVWLW 521
Db 340 --ILPVNKLISGSIATLANWLTSPAITMTNLTMTWSVGTSLSYMVSAFTIVFVWLW 397
QY 522 VPETKG 527
Db 398 VPETKG 403

RESULT 9
US-10-425-114-65425
; Sequence 65425, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 65425
; LENGTH: 318
; TYPE: PRT
```

```
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4764-010-D11_F1.1.pep
US-10-425-114-65425

Query Match
Best Local Similarity 43.9%; Score 1219; DB 15; Length 318;
Matches 246; Conservative 24; Mismatches 38; Indels 16; Gaps 5;

QY 9 PAGKKKKKQGLRREAVPGPASELRTRVNGGSGNRGGAGAGE-ESGSDHGVLRRLPL 67
Db 6 PTERKKNPNL-----AASWAARAGAMARSGRTSGGVEYSGSDHGVLRRLPL 54
QY 68 -NTGSWYR--MSSRQSSP-APGTSMAVLRSHVSAFLCTLIVALGPQFGFTSGFSSPT 123
Db 55 PNSGSWYRGMGSGRQSSLNAAAGTSSMAVLRSHVSAFLCTLIVALGPQFGFTSGFSSPT 114
QY 124 ODAMVRDNLNLSISFSAFGSLSNVGGVMGAIASQMAEYIGRKSLMTAAIPNIIGLWLA 183
Db 115 QASIRDLNLSISFSAFGSLSNVGGVMGAIASQMAEYIGRKSLMTAAIPNIIGLWLA 174
QY 184 SFAKDASFLYNGRLLEGFGVGIISYTPVPIAEISPNMRGALGSVNQLSVTFGLFAYL 243
Db 175 SFAKDSSFLYNGRLLEGFGVGIISYTPVPIAEISPNMRGALGSVNQLSVTFGLFAYL 234
QY 244 LGMFIPWRLAVIGALPCTMLIPGLFFIPESPRWLAKNLTDCETSLQVLRGFTDITT 303
Db 235 LGLFVWRLAVIGALPCTMLIPGLFFIPESPRWLAKNLTDCETSLQVLRGFTDITT 294
QY 304 EVNDIKRAVASSKRTTISFOELN 327
Db 295 ELNDIKRAVMSANKRATIRFQELN 318

RESULT 10
US-10-424-599-275268
; Sequence 275268, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 275268
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(361)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_90588C.1.pep
US-10-424-599-275268

Query Match
Best Local Similarity 43.1%; Score 1199; DB 15; Length 361;
Matches 226; Conservative 63; Mismatches 52; Indels 6; Gaps 1;

QY 187 KDASFLYNGRLLEGFGVGIISYTPVPIAEISPNMRGALGSVNQLSVTFGLFAYL 246
Db 4 QDSFLYNGRLLEGFGVGIISYTPVPIAEISPNMRGALGSVNQLSVTFGLFAYL 63
QY 247 FIPWRLAVIGALPCTMLIPGLFFIPESPRWLAKNLTDCETSLQVLRGFTDITT 306
Db 64 FVNRVLAIGLPCTMLIPGLFFIPESPRWLAKNLTDCETSLQVLRGFTDITT 123
```


QY 29 RPASELRTRVMGSGNRGGAGAGESSDHDGVLRRPRL-----NTGSW 72
Db 5 RAEDVEGAVAGAGGGGGG-----GEVTAPLLRLQHQKQGBEBKIQNDAGR 56
QY 73 YRMSRQSSFPAGPCTSSMAVLRSHVSAFLCTLVALGPQFGTSQFSSPTQDAMVRDLN 132
Db 57 LRGSARR-----KRRRV-----GYSAPTQSGIVDEVG 83
QY 133 LSISEFSAFSLNVGSGMAIASGQVAEYIGK-----GSLMIAAIPIHIGWLAI 183
Db 84 LSIQFALFSGVLITIGAMICAVTSGRLADPLGRMVCHEFRISQTKRISATCIFGWLSS 143
QY 184 SPKADASFLYMGRLLEGFGVGIISYTVVPIABISQNMRGALSGVNLQSVTFGIFLAYL 243
Db 144 HLAGKVMILYFGRILLFGSTGILSYVVPVFIABIAPKNLRGLGATSNQLLCSGSSATVI 203
QY 244 LGMPIRLLAVLALPCTMLIIFGLPIFIPESPRWLAKMNLTECETSQVLRGFEFTDIT 303
Db 204 IGAVALWRNLVIVGIVPCVLLTGLLFIPEPRWLANVGEKEFHASLOMLRGEDADVSE 263
QY 304 EVNDIKRAVASSKRTTISFOELNOKKYRTPLLLGIGLLVLQNLGSGINGVLFYASSIFKA 363
Db 264 EAVEKEYIESLRFPKARVODLFLRKNIVAVTVGVGLMIFQQLGGINGVGFYASSIFTS 323
QY 364 AGVNSDLATCSGAIQVLATGVTTWLLDRAGRILLIISTSGMTCLLAVSVVFFLKN 423
Db 324 AGPSGK-LGRTILICIIQIPITLFGAILMDKSGRRVILMWSASGTFGLGCTLTGISFYLK-- 380
QY 424 ISQDSNSYILTMISLVGIVSFVITPSFGMGAIPLWMMSEILPVSTKSLGSGTATLANWL 483
Db 381 -AQLFSEW-PELALTGILVIGVSGIGMGPFWVMSEIFSIDKAIKGGSLVTLVSWL 438
QY 484 TSFAITMTNLMTWSYGGTFLSGYVVSFAFTIVFVVLWVPETKG 527
Db 439 GSFAISYSPFLMDWSSAGTFPMSAASLITILFVVMVVPETKG 482

RESULT 14

US-10-310-154-725
; Sequence 725, Application US/1031010:54
; Publication No. US20030233670A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; APPLICANT: Chomet, Paul S.
; APPLICANT: Adams, Thomas H.
; APPLICANT: Ruff, Thomas G.
; APPLICANT: Agarwal, Ameeta K.
; APPLICANT: Ahrens, Jeffrey E.
; APPLICANT: Ball, James A.
; APPLICANT: Banu, G.
; APPLICANT: Bell, Erin
; APPLICANT: Boddupalli, Raghava
; APPLICANT: Deikman, Jill
; APPLICANT: Deng, Molian
; APPLICANT: Dong, Jinchao
; APPLICANT: Duff, Stephen M.
; APPLICANT: Galligan, Meghan M.
; APPLICANT: Hinchey, Brenda S.
; APPLICANT: Huang, Shihshieh
; APPLICANT: Johnson, G. Richard
; APPLICANT: Jung, Vincent
; APPLICANT: Kretzmer, Keith A.
; APPLICANT: Laccetti, Lucille B.
; APPLICANT: Lai, Chao-Qiang
; APPLICANT: Lee, Gary
; APPLICANT: Lin, Jie-Yi
; APPLICANT: Liu, Jingdong
; APPLICANT: Lu, Bin
; APPLICANT: Luethy, Michael M.
; APPLICANT: Lund, Adrian
; APPLICANT: Madson, Linda L.
; APPLICANT: Malloy, Kathleen A.
; APPLICANT: McKiel, Christine L.

; APPLICANT: Miller, Philip W.
; APPLICANT: Padmavathi, Manchikanti
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Start, William G.
; APPLICANT: Tennessee, Dan
; APPLICANT: Vidya, K.R.
; APPLICANT: Wang, Haiyun
; APPLICANT: Xin, Zhanquo
; APPLICANT: Xu, Nanfei
; APPLICANT: Yang, Chunzhi
; APPLICANT: Zeng, Xiaoping
; APPLICANT: Zhang, Qiang
; APPLICANT: Zhao, Yajuan
; APPLICANT: Zhou, Li
; TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
; FILE REFERENCE: 38-15(52796)B
; CURRENT APPLICATION NUMBER: US/10/310,154
; CURRENT FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: 60/337,358
; PRIOR FILING DATE: 2001-12-04
; NUMBER OF SEQ ID NOS: 736
; SEQ ID NO 725
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Glycine max
US-10-310-154-725

Query Match 32.1%; Score 891; DB 14; Length 481;
Best Local Similarity 37.2%; Pred. No. 1.2e-68;
Matches 188; Conservative 97; Mismatches 171; Indels 50; Gaps 4;

QY 48 AGAGESSDHDGVLRRPRLNTGSHYRMSRQSSPAPCTSSMAVLRSHVSAFLCTLIVA 107
Db 25 AAACKENGSD-----KSVNGSIGMVL-----LSTLVAV 53

QY 108 LGPIQFGFTSGFSSPTQDAMVRDLNLSISEFSAFGLSNVGMVGAIASGQVAEYIGRK 167
Db 54 CGSFTFGTCVGYSAFTQAARADLNLSEAFSMFGLSVTIGALGATSGRITDFGRKG 113

QY 168 SLMIAAIPIHIGWLAIISPAKASFLYMGRLLEGFGVGIISYTVVPIABISQNMRGAL 227
Db 114 AMRISTGFCITGWIADVFFSKGSYSLDFGFTGYGIGVISYVVPVYIAEIPAKNLRGGLA 173

QY 228 SYNQLSVTFGIFLAVLLGMFIPWRLAVIGALPCTMLIPGLFFIPESPRWLAKMNLTEDC 287
Db 174 TTNQLLIVTGSVSVLLGSGVINRELALAGLAVPCICLLVGLCFIPESPRWLAKVGREKEF 233

QY 288 ETSLOVLRGFETDITTEVNDIKRAVASSKRTTISFOELNOKKYRTPLLLGIGLLVLQNL 347
Db 234 QLALSRLRGKDADISDEAAEILDYIETLQSLPKTKLLDLFQSKYVHVSVVIGVLMACQOS 293

QY 348 SGINGVLFYASSIFKAAGVNTSDDLATCSLGAIOVLATGVTTWLLDRAGRILLIISTSGM 407
Db 294 VGINGIGFYTAIFVAAGLSSGKAGTIYACIQIPFTLIGAILMDKSGRRPLWVMSAAGT 353

QY 408 TLCLLAVSVVFFLKDNISSQDSNSYVILTMISLVGIVSVFTPSFGMGAIPLWMMSEILPV 467
Db 354 FLGCFVAAPAFFLKD-----QSLLPEWVPIAFAGVLIVIAAFPSIGLGSVPVIMSEIFPI 409

QY 468 SIKSLGSGTATLANWLTSFAITMTNLMTSVSGGTFSLSYVMVVSFAFTIVFVVLWVPETKG 527
Db 410 HLKGTAGSLVWLVAWLGAWVSYTTFNLSMSSPGTLFLYAGCSSLTILFVAKLVPETKG 469

QY 528 XNSRGDTIFVSLSIORQQLWPECLS 553
Db 470 -----KTLSEIQACIS 480

RESULT 15
US-10-424-599-275267
; Sequence 275267, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 13, 2004, 11:39:29 ; Search time 6383 Seconds

(without alignments)
4097.007 Million cell updates/sec

Title: US-10-051-909-36

Perfect score: 2779
Sequence: 1 PSSSSFRPAGKKKKKQ3.....TIFVSLSTQRQLWLPECL5 553

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2 1/USPTO.spool_h/US10051909/runat_13102004_123336_19906/app_query.fasta_1.1678
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADING=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051909@cgn 1 1 9091 @runat_13102004_123336_19906 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2133	76.8	1870	8 AK100792	AK100792 Oryza sat
2	2017.5	72.6	2266	8 AK065497	AK065497 Oryza sat
3	1821.5	65.5	1948	8 AK068594	AK068594 Oryza sat
4	1757.5	63.2	2020	8 BVU43629	U43629 Beta vulgar

RESULT 1	AK100792	1870 bp	mRNA	linear	PLN 24-JUL-2003
LOCUS	Oryza sativa (japonica cultivar-group) cDNA clone J023121D20, full insert sequence.				
DEFINITION	AK100792				
ACCESSION	AK100792.1				
VERSION	GI:32986001				
KEYWORDS	FLI CDNA; CAP trapper.				
SOURCE	Oryza sativa (japonica cultivar-group)				
ORGANISM	Oryza sativa (japonica cultivar-group)				
REFERENCE	1				
AUTHORS	The Rice Full-Length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-Length cDNA Project Team; Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Iazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,S., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Ootomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kuraaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN, Kawai,J., Carrincci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,				

ALIGNMENTS

5	1746.5	62.8	1557	8	AY093274	Arabidops
6	1746.5	62.8	1866	8	AY059848	Arabidops
7	1738.5	62.6	1695	8	AF412060	Arabidops
8	1734.5	62.4	1464	8	AY124845	Arabidops
9	1726.5	62.1	1914	8	AF462803	Arabidops
10	1710	61.5	1377	6	AX653119	Sequence
11	1693	60.9	1825	8	BT014234	Lycopersi
12	1577	56.7	1428	6	AX653120	Sequence
13	1453.5	52.3	204649	8	AC120986	Oryza sat
14	1137.5	40.9	120977	8	AC025808	Genomic s
15	1117.5	40.2	68041	8	AC025814	Arabidops
16	992.5	35.7	99587	8	AC007858	Oryza sat
17	982	35.3	156643	8	AC120988	Oryza sat
18	965.5	34.7	1699	8	LES278765	Lycopersi
19	954.5	34.3	2008	8	AK106658	Oryza sat
20	951.5	34.2	1741	8	AY084684	Arabidops
21	944.5	34.0	1778	8	AY064144	Arabidops
22	943.5	34.0	1435	8	BT000053	Arabidops
23	943.5	34.0	1686	8	AY120715	Arabidops
24	939	33.8	1449	8	BT000608	Arabidops
25	935.5	33.7	1747	8	AY087180	Arabidops
26	932	33.5	1806	8	BT008661	Arabidops
27	932	33.5	1847	6	ATH249967	Arabidops
28	920.5	33.1	1473	6	AX507636	Sequence
29	920.5	33.1	1678	8	D89051	Arabidops
30	899	32.3	1359	6	AX506932	Sequence
31	898.5	32.3	1420	8	AY091216	Arabidops
32	898.5	32.3	1660	8	AY063856	Arabidops
33	891.5	32.1	1755	8	AK064996	Oryza sat
34	878.5	31.6	1729	8	AF367260	Arabidops
35	877.5	31.6	1407	6	AX506596	Sequence
36	877.5	31.6	1407	6	AX651294	Sequence
37	877.5	31.6	1413	8	AY133547	Arabidops
38	870.5	31.3	1606	8	AY026255	Arabidops
39	866	31.2	1578	8	APH249968	Arabidops
40	858.5	30.9	1757	8	AY048207	Arabidops
41	858.5	30.9	1757	8	AY123991	Arabidops
42	853.5	30.7	1787	8	AY026254	Arabidops
43	845	30.4	1668	8	AY058056	Arabidops
44	843.5	30.4	1377	8	BT002699	Arabidops
45	837.5	30.1	1209	6	AX653566	Sequence


```

Qy 314 SerSerSerlyAArgThrThrIleSerPheGlnGlnLeuAsnGlnlyAslyTyrrArgThr 333
Db 999 TCAGCAAAACAAAGGACAAAGATCCGCTTTTTCAGAAATTAACAGAGAAATACCGCAC 1058
Qy 334 ProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyVal 353
Db 1059 CCCTTAATCTAGAAATGGCTTACTCTGTCTACTGACACAGCTAAGTGGATCATGGAATA 1118
Qy 354 LeuPheTyAlaSerSerIlePheIleAlaGlyValThrAsnSerAspLeuAlaThr 373
Db 1119 TTGTTTTATGCAGGTAGCATCTTCAAAAGCAGCAGGTCTCACAAACAGTGCATGGCTACA 1178
Qy 374 CysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrIleLeuLeuAspArg 393
Db 1179 TGTGCACCTTGGTGTCTTCAGGTTCTGCTACAGGAGTTCACACCTGGTTATTAGACACA 1238
Qy 394 AlaGlyArgArgIleLeuLeuIleSerThrSerGlyMetThrLeuCysLeuLeuAla 413
Db 1239 GCTGGCCGCGGATCTCTCTATCATCTCTTCTGCTGGGATGACTCTAAGCCCTCTTGCA 1298
Qy 414 ValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrrIle 433
Db 1299 GTTGTCTGTGTATTTTCTCAAGGATAGCATTTCAAGATTCTCACATGTACTACACC 1358
Qy 434 LeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMet 453
Db 1359 TTAAGTATGATCTCTCTGTGCTCTGTGCTTTTGTAAATCCCTTCCTTCGGTATG 1418
Qy 454 GlyAlaIleProThrLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGly 473
Db 1419 GGTGCCATTCATGATCATATGATGATGATGATGATGATGATGATGATGATGATGATG 1478
Qy 474 GlySerIleAlaThrLeuAlaAsnThrLeuThrSerPheAlaIleThrMetThrAsn 493
Db 1479 GGAAGCTTTGGAGCTGCGCACTGCTTACATCTCTTGGATATACATGACAGCAAC 1538
Qy 494 LeuMetLeuThrThrPheValGlyIleThrPheLeuSerTyrrMetValValSerAlaPhe 513
Db 1539 TTGATGCTTACGTGAGTGTGGAGGACCTTTGTGCTTACATGCTGCTGAGTGTCTTC 1598
Qy 514 ThrIleValPheValValLeuThrValProGluThrIleGly***AsnSerArgGlyAsp 533
Db 1599 ACCCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1657
Qy 534 ThrIlePheValSerLeuSerIle 541
Db 1658 ACATGCTCTCTCGCTGAGCCTT 1681

RESULT 2
AK065497
LOCUS Orzya sativa (japonica cultivar-group) cDNA clone:J013023P09, full
DEFINITION Orzya sativa (japonica cultivar-group) cDNA clone:J013023P09, full
insert sequence.
ACCESSION AK065497
VERSION AK065497.1 GI:32975515
KEYWORDS FLI_CDNA; CAP trapper.
SOURCE Orzya sativa (japonica cultivar-group)
ORGANISM Orzya sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Orzya.

REFERENCE
1. The Rice Full-length cDNA Consortium, National Institute of
Agricultural Sciences Rice Full-length cDNA Project Team,
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group, Ohtsuki, K., Ohtsuki, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,

```

```

Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusumegi, T., Ooka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN:
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashidume, W., Hayatsu, N., Imorani, K., Ishii, Y., Itoh, M.,
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Oka, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M., and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 (5631), 376-379 (2003)
22752273
12869764
2 (bases 1 to 2266)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kawaguchi, S., Kato, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,
Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
Osato, N., Ota, Y., Ohtsuki, K., Ohtsuki, K., Ohtsuki, K., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and
Yoshimura, A.
Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agricultural Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail: skikuchienias affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica
rice.
URL : http://cdna01.dna.affrc.go.jp/cdna/
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and
Yamamoto, M.
FAIS Genome Sequencing & Analysis Group: Ohtsuki, K., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
Yoshimura, A., Matsubara, K., and Murakami, K.
and Genome Science Research Group in Riken Genomic Sciences Center
Genome Exploration Laboratory in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Kanagawa, S., Kato, H., Kawai, J., Kouda, M.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,
Ota, Y., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A., and Hayashizaki, Y.
Location/Qualifiers
1. .2266
/organism="Orzya sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
FEATURES
source

```


528 *AsnSerArgGlyAspThrIlePheValSerLeuSerIleGln 542
1940 -AACTCTGGAGGAGATACAGTTTTCGTTCTGCTAAATTTGAG 1981

RESULT 3
AK068594

AK068594
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) cDNA clone:J03154G15, full insert sequence.
linear mRNA
1948 bp
PLN 24-JUL-2003

AK068594.1
AK068594.1
AK068594.1

KEYWORDS
FLI CDNA; CAP trapper.

SOURCE *Oryza sativa* (japonica cultivar-group)

ORGANISM *Oryza sativa* (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Equisetata; Magnoliophyta; Liliopsida; Poales; Poaceae; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartioideae; Oryzaceae; Oryza.

REFERENCE AUTHORS

The Rice Full-Length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-Length cDNA Project Team, Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yasaki, J., Ishikawa, M., Yamada, H., Ooka, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Otsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Otono, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Sugiyama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nariakawa, R., Kodama, A., Mizuno, K., Yokomizo, S., Miura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., KIKEN: Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osaco, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.

ETLII

JOURNAL
Science 301 (5631), 376-379 (2003)

MEDLINE 22752273

PUBMED 12869764

REFERENCE
AUTHORS

QY		1	ProSerGerSerSerPheArgProAlaGlyLysLysLysLysLysAsnGlnGly	20
			:::	:::
Db		17	CCCTCAATCAATCAAGCTTAAGCTTGGCGCGCAGCAGAGAAACGAGGAT	76
			:::	:::
QY		21	LeuArgGluAlaValProGlyArgProAlaSerGluLeuArgThrArgValMetGly	40
			:::	:::
Db		77	CGATCGCGTCACGCACGCCCGAT---CCC GCCGCGAGATGAGCTTCGCGCGCAGGAG	133
			:::	:::
QY		41	GlyGlySerAsnArgGlyGlyAlaGlyAlaGlyGluSerGlySerAspHisAspGly	60
			:::	:::
Db		134	AGCGCGCGCAGGATGGGGCGCGACGCGC-----TCGCGCTCGGAC-----	175
			:::	:::
QY		61	ValLeuArgArgProLeuLeuAsnThrGlySerTrpTyrArgMetSerSerArgGlnSer	80
			:::	:::
Db		176	-- --CTGGGAAGCCGTCTCTCCACACGGGAGCTGGTACAAGATGTCGTCCGCCGCCGCC	232
			:::	:::
QY		81	SerPheAlaProGlyThr-----SerSerMetAlaValLeuArgGluSerHisVal	97
			:::	:::
Db		233	GGCGCGGGATGGGTTCGGCGCTCGGCTCTCTCCGCCTACTCCCTCCGCGCACTCCTCCGTC	292
			:::	:::
			:::	:::

TITLE
JOURNAL

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan [E-mail: skikuchi@nias.affrc.go.jp].

COMMENT
This clone is one of the 28K full-length cDNA clones from japonica
Tel:81-29-838-7007, Fax:81-29-838-7007)

rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K.,

Nagata, T., Kawagashira, N., Dol, K., Kishimoto, N., Yazaki, J.,

229 ---GGAGATCTCGCGAAACCAATTTTTCATACAGGAAGTTGGTATAGAAATGGGTTCAGA 285
 79 GlnSerSerPheAlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSer 98
 286 CAATCAAGTTTG-----ATGGGTCTCTCTCAAGTTATTTCAGAAAGCTCCATCTCT 336
 99 AlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGly 118
 337 GTCCTCTGTTGTTCTTATTGTTGCTTGGTCTTATTCAGTTTGGTTTACTGCTGTT 396
 119 PheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPhe 138
 397 TATTCACTCTCAACTCAATCAGCAATACAAATGAAGCTTGGTCTTCTGTTGAGAGATC 456
 139 SerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGln 158
 457 TCATGGTTTGGTCTCTATCCAAATGTTGGTGCCATGTTGGGCAATTGCAAGTGTCAA 516
 159 MetAlaGluTrpIleGlyArgGlyGlySerLeuMetIleAlaIleProAsnIleIle 178
 517 ATATCTGAGTATATTGGAAGAAAGGGTCTTTGATGATTGCTGATCCCAACATAAAT 576
 179 GlyTrpLeuAlaIleSerPheAlaValAspAlaSerPheLeuTrpMetGlyArgGluLeu 198
 577 GGATGGCTTGCTATATCATTTGCCAAAGATCTTCATTTCTGATATGGGAAGATTTG 636
 199 GluGlyPheGlyValGlyIleIleSerTrpThrValProValTrpIleAlaGluIleSer 218
 637 GAAGGTTTGGCGTGGGAATAATATCTTACACGGTCTCTGTGTACATTTCTGAAATAGCA 696
 219 ProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIle 238
 697 CCACAGAAATTTGAGGGGAGCTTTGGGTTCAGTAACACAGCTATCCGTTTACAATCGGTATA 756
 239 PheLeuAlaTrpLeuLeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAla 258
 757 ATGTCATCGTATATGCTGGGCTTTTTCCTCGGTGGAGATACTTGCAGTTTATAGGAAT 816
 259 LeuProCysThrMetLeuIleProGlyLeuPheIleProGluLeuSerProArgTrpLeu 278
 817 TTACCATGCATATATTGATACCTGGCTTATTTCTTATACCGGAATCTCTAGTGTGGTTG 876
 279 AlalySerMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGlu 298
 877 CGGAATATGGTATGATGGAAGAAATTTGAGTTTCTTGTGCAAGTTCTCCGGGGCTTTGAC 936
 299 ThrAspIleThrThrGluValAsnAspIleGlySerAlaValAlaSerSerSerLysArg 318
 937 ACTGATATATCTTGAAGTTAATGAAATCAAGAGATCTGTGCTTCAAGTAGCAAAAGA 996
 319 ThrThrIleSerPheGlnGluLeuAsnGlnLysLysTrpArgThrProLeuLeuLeuGly 338
 997 ACCACCATAGATTTGCAGAACTACGACAAAGAGATATTGGTTACCTTGTATGATAGGG 1056
 339 IleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTrpAlaSer 358
 1057 AATGGTTTACTTATTGCAACAACCTCAGTGGATCAATGGTGTCTTCTTCTTCTTCTAGT 1116
 359 SerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAla 378
 1117 ACCATCTTTAAAGAGGCTGGGGTCCATCAAGCAACGCTGCAACATTTGGCCCTTGGTCT 1176
 379 IleGlnValLeuAlaThrGlyValThrTrpLeuLeuAspArgAlaGlyArgArgIle 398
 1177 GTTCAGGTTATCGCACTTGTGACTACCTGGTGGTGGATTAATCAGACGACGAGACTT 1236
 399 LeuLeuIleIleSerThrSerGlyMetThrLeuLeuCysLeuLeuAlaValSerValPhe 418
 1237 CTCTCAATGTCCTCTCTCTGGGATGACATTGATCTTCTTCTGATGGTGGCAATGCTCCTTC 1296
 419 PheLeuLysAspAsnIleSerClnAspSerAsnSerTrpTrpIleLeuThrMetIleSer 438
 1297 TTTTAAAGGAATGGTATACAGCAATCTTCATGGTACAGTGTATTCTTCATCTCTTCA 1356

439 LeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrp 458
 1357 GTTGTGGAGTTGTGGCTATGTTGGTTACTTCTTCTCATTAGGAATTGGAGCAATTCGGTGG 1416
 459 LeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThr 478
 1417 ATTAAATGTCAGAGATTCTTCCAATTAACATCAAGGGCTGGCTGGAGATGATTCTACA 1476
 479 LeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrAsnLeuMetLeuThrTrp 498
 1477 CTTCGCACTGGTTGTAGCATGGATTGGCAATGACTGCAACATATATGTTGTCATGG 1536
 499 SerValGlyGlyThrPheLeuSerTrpMetValValSerAlaPheThrIleValPheVal 518
 1537 AATAGTGGAGGTACTTCTTCTTATTTACATGGTTGTTGTCTTCTTCTTCTGCTTCTG 1596
 519 ValLeuTrpValProGluThrLysGly***AsnSerArgGlyAspThrIlePheValSer 538
 1597 GTCATTTGGTTCCTCCGACAGCAAGGGGAG-AACATTGGAGGAGATTTCAGT---GTCCTT 1652
 539 LeuSerIleGlnArgGlnLeuGlnTrpLeuProGluCys 551
 1653 CAGAAGATAAAGAAATTTTCTGATAACTTGTCCATGGTGT 1691
 RESULT 5
 AY093274 1557 bp mRNA linear PLN 21-APR-2002
 LOCUS Arabidopsis thaliana similar to integral membrane protein
 DEFINITION (Atg19450) mRNA, complete cds.
 ACCESSION AY093274.1 GI:20259851
 VERSION AY093274.1
 KEYWORDS Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 1557)
 Southwick,A., Karlin-Neumann,G., Nguyen,M., Lam,B., Miranda,M.,
 Palm,C.J., Bowser,L., Jones,T., Bann,J., Carninci,P., Chen,H.,
 Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
 Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
 Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
 Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
 Direct Submission
 Submitted (26-MAR-2002) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA
 e-mail for correspondence: arab@sequence.stanford.edu
 RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN
 Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
 Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
 Hayashizaki,Y. and Shinozaki,K.
 The Salk, Stanford, PGEC (SSP) Consortium members carried out the
 sequencing and annotation of the RAFL cDNAs: Southwick,A.,
 Nguyen,M., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J.,
 Bowser,L., Jones,T., Bann,J., Chen,H., Cheuk,R., Chung,M.K.,
 Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P.,
 Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.
 Southwick,A., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed
 equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.
 (SSP/Stanford) contributed equally to this work as PIs.
 FEATURES
 source
 1. 1557
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="1"

/clone="U16352"
 /ecotype="Columbia"
 /note="This clone is in pENTR/SD-dtopo This is a cloned
 PCR product using RIKEN clone RAF107-13-K02 (AY059848) as
 a template"
 1..1557
 /gene="Atig19450"
 /note="synonym: F18014.22"
 1..1467
 /gene="Atig19450"
 /codon_start=1
 /product="similar to integral membrane protein"
 /protein_id="AA13273.1"
 /db_xref="GI:20259852"
 /translation="MSPRDDNTTEGRNDRPFLHTGSWYRMGRSSMLSESSQVIRD
 SSVLACVIALGPIQFGFTCYSSPTQQAIAITKDLGLTVSEYVFGSLSNVGMVG
 AIAQSGAEVYVGRKGSIMIAIPNIIGWLSISPAKDTSPLYMGRLLLEGFGVGIISVTV
 PVYIAETAPOTMRGALGVNQLSVITIGIMLAYLLGLFVPRILAILVGLVLPCTLLIPGL
 FFIPEPRWLAKGLTDDFTSLQVLEGFETDITVNEIKRSVASSSSKSAVRFDVL
 KRRYIFPLMWGGLLQQLGGINGVLFSSIFISAGVTSNNATFVGWVQVAVT
 GIATVLDKAGRRLLMISGIMTISLIVAVAFYLFKRFVPSDNNINILSVSVVGV
 VAMVISGLGPGPIMLIMSEIILFVNKIGLAGSIATLLANFVSLVMTANMLLAWSS
 GGTFTLYALVCGFTVVFVSLVWVPETKLTLEIQAIFR"
 ORIGIN

Alignment Scores:
 Pred. No.: 1,98e-121 Length: 1557
 Score: 1746.50 Matches: 338
 Percent Similarity: 83.40% Conservative: 69
 Best Local Similarity: 69.26% Mismatches: 74
 Query Match: 62.85% Indels: 8
 DB: 8 Gaps: 2

US-10-051-909-36 (1-553) x AY093274 (1-1557)

QY 52 GluGluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAenThrGlySer 71
 DB 25 GAAGAGGGAAGAAATGAT-----CTTCGGGACCGTCTTACACAGCAAGT 72
 QY 72 TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal 91
 DB 73 TGGTATCGGATGGGTTTCGAGACAATCAGATG-----TTGGAAATCGTCTCAAGTT 123
 QY 92 LeuArgGluSerHisValSerAlaPheLeuLeuCysThrLeuLeuValAlaLeuGlyProIle 111
 DB 124 ATTCGAGACAGCTCAATTTCTGTAGCTTGTTGTTTGTATGTTCTCTTGCTGCTAT 183
 QY 112 GlnPheGlyPheThrSerGlyPheSerProThrGlnAspAlaMetValArgAspLeu 131
 DB 184 CAATTCGGATTCACCTTGCTGTTATCTCTCAACCTCAAGCTCCCATTAAGGATCTT 243
 QY 132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal 151
 DB 244 GGTTRACTGTATCAGAGTACTCTGTGTTGGGTCCTATCCATGTGGTGCTATGTT 303
 QY 152 GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIle 171
 DB 304 GGAGCAATTCGAGTGGTTCAGATGCTGTAATAGCTTGGAGGAAAGGCTCTCTGATGATT 363
 QY 172 AlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPhe 191
 DB 364 GCTGCAATTCCTCAATATATGATGAGGCTTTCGATATCATTTGCAAAAGATCTCTTTT 423
 QY 192 LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro 211
 DB 424 CTTTATACATCGAAGATGTTAGAAGGCTTCGAGTGGGATATCTCTATACCGTTCCTCC 483
 QY 212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyValAlaLeuGlySerValaenGln 231
 DB 484 GTATATAGCTGAGATCGCTCCACAGACCATCAGAGGAGCCCTTAGGTCAGTTAACCCAG 543
 QY 232 LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArg 251

DB 544 CTTTCTGTAAACAATTGGGATAAATGCTGGCGTATTTACTCGGTCTCTTTGTTTCCATGGAGA 603
 QY 252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuLeuIleProGlyLeuPhePheIle 271
 DB 604 ATCTCTGAGTTTGGGAGTATTGCCATGTACATTTGTTGATACCGGGTCTTTTTCATT 663
 QY 272 ProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu 291
 DB 664 CCGTGAATCTCTCGTGGTGGTTCGCAAGATGGGTTTGACAGATGATTTTCGAACATTCATT 723
 QY 292 GlnValLeuArgGlyPheGluThrAspIleThrGluValAsnAspIleLysArgAla 311
 DB 724 CAAATCTCTCGTGGATTTGAGACTGATATACCGTTGAGGTTATGAATCAAGAGATCT 783
 QY 312 ValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysTyr 331
 DB 784 GTGGCATCATCTAGCAAAACGTTCTCGAGTTCGGTTGTAGACCTCAAGCGCAGAGATAC 843
 QY 332 ArgThrProLeuLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleasn 351
 DB 844 TATTTCCCACTGATGGTGGTATAGGCTGCTTCGACTTCAACAACTTGGAGGAATCAAT 903
 QY 352 GlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeu 371
 DB 904 GGTGTCCTTGTCTATTTCGAGTACAATTTTGAATCTCGAGGGGTGACATCAAGTAATGTG 963
 QY 372 AlaThrCysSerSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrLeuLeu 391
 DB 964 GCGACATTTGAGATTGGCGTGTGTCAGGTAGTGGGATGGGATAGGACATGTTGGTG 1023
 QY 392 AspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeu 411
 DB 1024 GATAAAGCAGGTGCTCGGCTTCTACTCATGATCTCTTATCGGATGACGATAAGCCTC 1083
 QY 412 LeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyr 431
 DB 1084 GTGATGTGAGCAGTTGCATTTTACCTTAAGAAATTTGTATCACCTGATTTCCAAATCATAC 1143
 QY 432 TyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPhe 451
 DB 1144 AACATCTTACATCATGTTTCCGTAGTGGAGTGTGGCTATGTTATTTCTGTCTCTA 1203
 QY 452 GlyMetGlyAlaIleProTrpLeuMetSerGluIleLeuProValSerIleLysSer 471
 DB 1204 GGAATGGACCAATTCATGCTGATGATGCTGAGATTTCTCCAGTCAATATATAAGGGT 1263
 QY 472 LeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThr 491
 DB 1264 TTAGCCGGAAGTATAGCAACTTTGTTAACTGGTTCGTGTCATGGTTAGTGACAACTACT 1323
 QY 492 ThrAsnLeuMetLeuThrTrpSerValGlyThrPheLeuSerTyrMetValValSer 511
 DB 1324 GCATAATATGCTGTAGCATGAGCAGTGGGGAACCTTCACTCTACGCTTGGTTGT 1383
 QY 512 AlaPheThrIleValPheValValLeuTrpValProGluThrLysGly***AsnSerArg 531
 DB 1384 GGATTCAGTGGGTTTGTGAGTCTTGGGTTCTTGGAGTCTTGGAGCTAAAGGAAA-AACGCTTGA 1442
 QY 532 GlyAspThrIlePheValSerLeu 539
 DB 1443 AGAGATCCCAAGCTTTGTCAGATG 1466

RESULT 6

AY059848 1866 bp mRNA linear PLN 04-NOV-2001
 LOCUS Arabidopsis thaliana similar to integral membrane protein
 DEFINITION (At1g19450; F18014.22) mRNA, complete cds.
 ACCESSION AY059848 GI:16648956
 VERSION AY059848.1
 KEYWORDS FLI_CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
AUTHORS

1 (Bases 1 to 1866)
Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M.,
Palm,C.J., Bowser,L., Jones,T., Bath,J., Carninci,P., Chen,H.,
Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (22-Oct-2001) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
e-mail for correspondence: arabseq@sequence.stanford.edu

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN
Arabidopsis Full-length cDNA': Seki,M., Narusaka,M., Ishida,J.,
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGC (SSP) Consortium members carried out the
sequencing and annotation of the RAPL cDNAs: Nguyen,M.,
Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J.,
Bowser,L., Jones,T., Bath,J., Chen,H., Cheuk,R., Chung,M.K.,
Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P.,
Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Nguyen,M. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed
equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.
(SSP/Stanford) contributed equally to this work as PIs.

FEATURES
source

Location/Qualifiers
1..1866
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAPL07-13-K02"
/note="This clone is in pBluescript"

gene

1..1866

CDS

249..1715
/gene="Atig19450; F18014.22"
/gene="Atig19450; F18014.22"
/codon_start
/product="similar to integral membrane protein"
/protein_id="AAL24330.1"
/db_xref="GI:16648957"
/translation="MSFRDNDTEGRNDRPFLHTGSWMRSGROSSMLESQVIRD
SSIVLACVLIVLGPQFGFTGCGSPQAAITKDLGLTVSEYSPGSLSNVGMVVG
ALASQIAEYVGRKSLMIAIPNIGWLSIPAKDTSFLYMGRLLEGFGVGIISVTY
PVYIARAPQMRGALGVNCLSVTIGIMLAYLLGLFVPMELAVLGLVPLTLLPGL
FFIPERFLAMGLDDEFTSLQVLRGFTDITVEVNEIKRSVASSKRSVRFDVL
KRRYYFPLMVGILLALQQLGGINGVLFFYSITFESAGVTSNNVATFGVVQVAVT
GIATWLDKAGRRLLMISSTIGTISLIVAVAFYLFKRVSPDSNNYINLSVSVGV
VAMVISCGLGMPFPMIMSEILFVNKIGLAGSIATLLNFWFVSLVMTANMLLAWSS
GGTFTLVLCVCGFTWVFVSLVWPBTKGKLTLEEQALFR"

ORIGIN

Alignment Scores:

Pred. No.: 2,55e-121 Length: 1866
Score: 1746.50 Matches: 338
Percent Similarity: 83.40% Conservative: 69
Best Local Similarity: 69.26% Mismatches: 74
Query Match: 62.85% Indels: 8
DB: Gaps: 2

US-10-051-909-36 (1-553) x AY059848 (1-1866)

QY

52 GluGluserGlySerAspHisGlyValLeuArgProLeuLeuAsnThrGlySer 71

DB

273 GAAGAGGGAAGAATGAT-----CTTCGGCGACCGTCTTACACACAGGAAGT 320

QY

72 TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal 91

321 TGGTATCGGATGGGTTTCGAGACAACTAGTAGT-----TTGGAATCGTCTCAAGTT 371
92 LeuArgGluSerHisValSerAlaPheLeuCysThrLeuValAlaLeuGlyProLeu 111
372 ATTCGAGACAGCTCAATTTCTGCTTCTAGCTTGTTTGAATGTTGCTCTTGCTGCTTAT 431
112 GlnPheGlyPheThrSerGlyPheSerProThrGlnAspAlaMetValArgAspLeu 131
432 CAATTTCGGATTTCATTTGGTTATTCCTCACCACCTCAAGCTGCCATCTACTAAGGATCTT 491
132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyMetVal 151
492 GCTTTAATCTGATCAGAGTACTCTGTTGGTCTCTATCCAAATGGTGGTCTAAGTT 551
152 GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIle 171
552 GAGCAATTCGAGTGGTCAGATTCCTGATACGTTGCGAGCAAGAGGTCTCTGATGATT 611
172 AlaAlaIleProAsnIleGlyTyrLeuAlaIleSerPheAlaLysAspAlaSerPhe 191
612 GCTGCAATTCCTCAATATTATTGGATGGCTTTGATATCATTTGCAAAAGATACTTCTTT 671
192 LeuTyrMetGlyArgLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro 211
672 CTTTACATCGGAGAGATTGTTAGAGGCTTCGAGTTGGATATCTCTATACGTTCC 731
212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231
732 GTATATATAGCTGAGATCGCTCCACAGACCATGAGAGGAGCCTTAGGTTACGTTAACCCAG 791
232 LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTrpArg 251
792 CTTTCTGTAACTTGGGATAATGCTGGGATATTACTCGGTCTCTTTGTTCCATGGAGA 851
252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIle 271
852 ATTCITTCAGTTTTCGGAGTATTGCCATGATCATTTGTTGATACCGGTCTTTTTCATT 911
272 ProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu 291
912 CTGTAATCTCTCGGTGGTGGCAAGATGGGTTTGACAGATGATTTGAAACTTCATTG 971
292 GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAla 311
972 CAACTTCTTCGTGGATTGAGACTGATATACCGTTAGGTTAATGAAATCAAGAGATCT 1031
312 ValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyr 331
1032 GTGGCATCATCTAGCAAAACGTTCTGCAGTTCGGTTTGTAGACCTCAAGCGCAGGAGATAC 1091
332 ArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsn 351
1092 TATTTCCCACTGATGTTGGTATAGGCTGCTTGACCTCAACAACTTGGAGGAATCAAT 1151
352 GlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeu 371
1152 GGTGCTCTTCTTCTATTTCGAGTACAATTTTGAATCTGCGAGGGGTGACATCAAGTAATGTG 1211
372 AlaThrCysSerLeuGlyValAlaIleGlnValLeuAlaThrGlyValThrThrProLeuLeu 391
1212 GCGACATTTGGAGTTGGCGCTTGTTCAGGTAGTGGGACCTGGGATACGACATCGTTGGTG 1271
392 AspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeu 411
1272 GATAAGCAGGTCGTGGGCTTCTACTCATGATCTCTTCTATCGGAATGACGATCAAGCCTC 1331
412 LeuAlaValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyr 431
1332 GTGATGTGACAGTTGCAATTTTACCTTAAGGAATTGATCATCCTGATTCACATGATGAC 1391
432 TyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPhe 451

Db 1392 AACATTCTAACAGCGTTTCCTCGTAGTTGGAGTTTGGCTATGGTTATTTCTTCTCTCTCTA 1451

Qy 452 GlyMetGlyAlaIleProTrpLeuMetSerGluIleLeuProValSerIleLysSer 471

Db 1452 GGAATGGGACCAATCCATGCTGATATGCTGAGATTCTGCCAGTGAATATAGGGT 1511

Qy 472 LeuGlyGlySerIleAlaIleLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThr 491

Db 1512 TTACCCGGAAGTATAGCAACTTTCTTAACTGGTTCGTGTCATGCTAGTACACATCACT 1571

Qy 492 ThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTrpMetValValSer 511

Db 1572 GCAATATGCTGTAGCATGGAGCAGTGGGGACATTTCTACTCTACGCTTTGGTTGT 1631

Qy 512 AlaPheThrIleValPheValValLeuTrpValProGluThrIleGly***AsnSerArg 531

Db 1632 CGAATCACTGTGTTTGTGAGTCTTTGGGTTCTCTGAGACTAAAGAAA-AACGCTTGA 1690

Qy 532 GlyAspThrIlePheValSerLeu 539

Db 1691 AGAGATCAAGCTTTGTTCAGATG 1714

RESULT 7

LOCUS AF412060 1695 bp mRNA linear PLN 24-SEP-2001

DEFINITION Arabidopsis thaliana At1g75220/F22H5_6 mRNA, complete cds.

ACCESSION AF412060

VERSION AF412060.1 GI:15724239

KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1. (bases 1 to 1695)

Chen, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Ban, J., Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

Arabidopsis cDNA clones

Unpublished

2. (bases 1 to 1695)

Chen, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Ban, J., Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

Direct Submission

Submitted (21-AUG-2001) Salk Institute Genomic Analysis Laboratory (Signal), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGECC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Chen, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Shinn, P., Ban, J., Bowser, L., Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Chen, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

Location/Qualifiers

1..1695

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="1"

/clone="RAF106-75-L22 (R13497)"

/ecotype="Columbia"

1..11

12..1475

/note="putative integral membrane protein"

/codon_start=1

/product="At1g75220/F22H5_6"

/protein_id="AA106513.1"

/db_xref="GI:15724240"

/translation="MSFRDNEAEANDLRPFHTGSMYWGSPQSSMMGSSOVIDRS
SISVACVLIIVAGLPIDFGTCGYSSTQAIKDKLGLTVSEYVFSLSNVGAVGA
IASQAIYIGRKGLSMALPNIIGWLCISFAKDTFLXGRLLGFGVGLISYIVP
VYIAIAPQNNRGGLSVNQLSVITIGIMLAYLGLFVFWRLAVGLPCTLLPGLF
PIPSRPLAKMGMTDEFETSLQVLRGFETITVEVNEIKRVSASTKRVNTVFDLK
RRVYFPLMVGIGLLVQLQGLNGVLFYSSTIFESAGVTSSNAATFGVAIQVATA
ISTWLDKAGRPRLITISSVGMTISLVIAAAYLKEFVSPSDMYSLWSILSVGVV
AMVFFSLGNGPIPLIMSEILPVNKLGLAGSIATLANWPFSLMTITANLLAWSSG
GTFTYGLVLCATVTVFVTLWVPEYKTKLEELQLPR"

1476..1695

3'UTR

ORIGIN

Alignment Scores:

Pred. No.: 8 87e-121 Length: 1695

Score: 1738.50 Matches: 343

Percent Similarity: 82.47% Conservative: 71

Best Local Similarity: 68.33% Mismatches: 81

Query Match: 62.56% Indels: 8

DB: 8 Gaps: 2

US-10-051-909-36 (1-553) x AF412060 (1-1695)

Qy 52 GluGluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySer 71

Db 33 GAGGAGGCTAGGAATGAT-----TTGCGTCGCGCGTTTATTCACACCGAAGC 80

Qy 72 TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal 91

Db 81 TGGTACCGGATGGTGGTTCGAGACATCGAGTATG-----ATGGGTTCTCTCAAGTT 131

Qy 92 LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIle 111

Db 132 ATTGAGACAGCTCTATCT 191

Qy 112 GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu 131

Db 192 CAATTCGGATTCATTTGGTATTATTCATTCCTCACTCACTCACTCACTCACTCACTCACT 251

Qy 132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyMetVal 151

Db 252 GGTTTAACAGTATCTGAGTACTCTGTGTGGTTCTCTCTCTCTCTCTCTCTCTCTCTCT 311

Qy 152 GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyValGlyGlySerLeuMetIle 171

Db 312 GGTGCAATTCAGTGGTTCAGATTGCAGATATACATTGGAGAAAGGCTCTCTATGATT 371

Qy 172 AlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPhe 191

Db 372 GCTCAATTCCTCAATATCATTTGGATGCTTTGGCATATCATTTGCCAAAGATACATCTCTTT 431

Qy 192 LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro 211

Db 432 CTTTACATGGGAAGACTTTTAGAAGGCTTTTGGGCTTTGGGATCATCTCTTACACTGTACCT 491

212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231
 492 GTATATATCGTGTAGATAGTCCACAAATATGAGAGGAGATGGGTTCGGTAAACCAG 551
 232 LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrPArg 251
 552 CTTTCGTGTACAAATGCTTGGCGTATTACTTGGCCTCTTTGTTCATGGAGA 611
 252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIle 271
 612 ATTTCTTCAGTTCCTGGGATATTGCGGTATACATATTATGATACAGGCTCTCTTTTCATC 671
 272 ProGluSerProArgTyrLeuAlaLeuMetAsnLeuThrGluAspCysGluThrSerLeu 291
 672 CCCGAATCCCTCGCTCGCTGGCAAAATGGGTATGACATGAATTTGAACTTCATTA 731
 292 GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAla 311
 732 CAAGTCTTCCGAGATTGAGACTGATATTACCGTTGAAGTTAATGAATCAAGAGATCT 791
 312 ValAlaSerSerIleValThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyr 331
 792 GTGCATCATCTCAAAAGCGAATATACAGTTCGGTTCCTGATCTCAAAACGACGATAC 851
 332 ArgThrProLeuLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsn 351
 852 TATTCCCACTTATGTTGGTATAGGTTGCTTCTACTTCAACAACTTGGTGAATTAAT 911
 352 GlyValLeuPheTyrAlaSerSerIlePheLysAlaGlyValThrAsnSerAspLeu 371
 912 GGTCTCTATTCTATCCAGTACATATTTGAATCTCGAGGATTTACATCGAGTAACGCA 971
 372 AlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeu 391
 972 GCAACATTTGGGTCGGGCTATTTCAGTAGTACGACTGCAATATCCATCTGTTGGTG 1031
 392 AspArgAlaGlyArgAlaGlyLeuLeuIleSerThrSerGlyMetThrLeuCysLeu 411
 1032 GACAAAGCAGTCTCGGCTCTCTTACTATCTCTCGTTCGGTGGGATGACGATTCGCT 1091
 412 LeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyr 431
 1092 GTAAATGTTCAGCTGCTTCTATCTTAAGGAATTTGTCTCTGATTCAGACATGATAC 1151
 432 TyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPhe 451
 1152 AGTTGGCTGACCATATTGTCAGTAGTGGAGTTCGGCAATGGTTGCTTTTCTCATTTG 1211
 452 GlyMetGlyAlaIleProThrLeuMetMetSerGluIleLeuProValSerIleLysSer 471
 1212 GGAATGGGACCAATACCGTGGCTCATTATGCTCAGATCCCTCTGTGAACATAAAGGT 1271
 472 LeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThr 491
 1272 TTAGTGGAGATATTGCAACTAGCCAAATGGTTCTTTCTTCTGTTGATCACCATGACA 1331
 492 ThrAsnLeuMetLeuThrTrpSerValGlyThrPheLeuSerTyrMetValSer 511
 1332 GCAAAATTTGCTGTAGCTGAGCAGTGGAGAACTTCACTCTGTATGGATGCTTTGT 1391
 512 AlaPheThrIleValPheValValLeuTrpValProGluThrLysGly***AsnSerArg 531
 1392 CATTTCAGTGGTGTTCGTGACTTATGGTTCCTGAGACAAAGGCAA-AACTCTTGA 1450
 532 GlyAspThrIlePheValSerLeuSerIleGlnArgGlnLeuGlnTrpLeuProGluCys 551
 1451 AGAATTCATCTTGTTCAGATGAACAAATGAACAACTTCATCTTCTTGTACCCCTCT 1510
 552 LeuSer 553
 1511 CTCTCC 1516

RESULT 8

AV124845 1464 bp mRNA linear PLN 06-JUL-2002
 LOCUS Arabidopsis thaliana At1g75220/F22H5_6 mRNA, complete cds.
 DEFINITION AY124845
 ACCESSION AY124845
 VERSION AY124845.1 GI:21700860
 KEYWORDS FLI CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 1464)
 Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Banh,J., Bowser,L.,
 Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
 Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
 Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
 Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
 Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,
 Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.
 and Ecker,J.R.
 TITLE Arabidopsis ORF clones
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1464)
 Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Banh,J., Bowser,L.,
 Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
 Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
 Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
 Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
 Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,
 Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.
 and Ecker,J.R.
 TITLE Direct Submission
 JOURNAL Submitted (17-JUN-2002) Salk Institute Genomic Analysis Laboratory
 (SIGNAL), Plant Biology Laboratory, The Salk Institute for
 Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
 USA
 COMMENT The two base pair differences do not change the amino acid sequence
 of the protein.
 RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL cDNAs (RAFL cDNA : RIKEN
 Arabidopsis Full-length cDNA) : Seki,M., Narusaka,M., Ishida,J.,
 Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
 Hayashizaki,Y. and Shinozaki,K.
 The Salk, Stanford, PGEC (SSP) Consortium members constructed and
 sequenced the PUNI (ORF) clones using the RAFL cDNAs: Cheuk,R.,
 Chen,H., Kim,C.J., Shinn,P., Banh,J., Bowser,L., Chan,M.M.,
 Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Jones,T.,
 Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M.,
 Onodera,C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C.,
 Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Davis,R.W.,
 Theologis,A., and Ecker,J.R.
 Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
 this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
 contributed equally to this work as PIs.
 FEATURES
 source Location/Qualifiers
 1. .1464
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="1"
 /clone="U13497"
 /scotype="Columbia"
 /note="This clone is in PUNI 51"
 1. .1464
 /note="putative integral membrane protein"
 /codon_start=1
 /product="At1g75220/F22H5_6"
 /protein_id="AAW0554.1"
 /db_xref="GI:21700861"
 /translation="MSFRDNEEARNDLRPFHTGSWYRMGROSSVMGSSQVIRDS
 SISVLACVLVALGPIQFGTICGYSPTQAATKDLGLTVSEYVFGSLNVMGVGA

IASGOIAEYIGRKGSLMIAAIPNLIIGWLCISFRAKDTSLYMGRLLEGFGVGIISYTPV
 VVIEAIPQNMREGGLSVNOLSVITIGIMLAVLGLFVPEILAVLIGLIPCTLLIPGLF
 FIPSPRLAKMGMDFEESLQVIRGFEDTIVENEIKRSVASSTKNTVRFVDFLK
 RRRYFLMGVIGLVLQOUGGNGVLFVIFESAGVTSNNAATFVGVAIOVAVATA
 ISTWLKAGRRLLLTSSVGMITSLIVAAAFYKFEFVSPDSDMYSLISVGVV
 AMVFFSLGMPILPMLJMSILPVNIKGLAGSIATLANWFFSLIITWNTANLLAWSSG
 GTFLYGLVCAFTVFTVTLVFPETKGTLEBLOSLFR

misc_difference 1463

/note="compared to genome"

/replace="a"

misc_difference 1464

/note="compared to genome"

/replace="g"

ORIGIN

Alignment Scores:
 Pred. No.: 1,44e-120 Length: 1464
 Score: 1734.50 Matches: 338
 Percent Similarity: 83.40% Conservative: 69
 Best Local Similarity: 69.26% Mismatches: 74
 Query Match: 62.41% Indels: 8
 DB: 8 Gaps: 2

US-10-051-909-36 (1-553) x AY124845 (1-1464)

QY 52 GluGluSerGlySerAspHisAspValLeuArgProLeuLeuAsnThrGlySer 71
 DB 22 GAGGAGGCTAGGAATGAT-----TTGGTCGGCGCTTATTACACCGGAAGC 69
 QY 72 TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal 91
 DB 70 TGGTACCGGATGGGTCGAGACATCAGATG-----ATGGGTCCTCTCAAGTT 120
 QY 92 LeuArgGluSerHisValSerAlaPheLeuCysThrLeuLeuValAlaLeuGlyProIle 111
 DB 121 ATTCGAGACAGCTCTATCTCTGTTCTTGTGATGTTTGTGCTCTTGGCCCTATT 180
 QY 112 GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu 131
 DB 181 CAATTCGGATTCACCTTGTGTTATTCATCCCAACTCAAGTCGAATCAACGAGATCTC 240
 QY 132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal 151
 DB 241 GGTTTACAGTACTCGATCTCTGTTTGTCTCTTCCATGTGGGTGCTATGTT 300
 QY 152 GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgGlySerLeuMetIle 171
 DB 301 GGTGCAATTCGCAATGTCAGATTCAGAAATACATTCGACGAAAGGGTCTCTTATGAT 360
 QY 172 AlaAlaIleProAsnIleIleGlyTyrLeuAlaIleSerPheAlaLysAspAlaSerPhe 191
 DB 361 CTGTCGAATTCGAATATCATATGATGATGCTTGGCATATCATTTGCAAAAGATCTCTTT 420
 QY 192 LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro 211
 DB 421 CTTTATACATGGAAGACTTTTAGAAGGCTTTGGCGTTGGGATCATCTCTTACACTGTACCT 480
 QY 212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyValAlaLeuGlySerValLeuGln 231
 DB 481 GTATATATCGCTGAGATGATGTCACAAATATAGAGGAGGATGGGTTCGGTTAACAG 540
 QY 232 LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProThrArg 251
 DB 541 CTTTCTGTGCAATTCGAATTAATGTTGGGTATTTACTTGGCCCTCTTCTTCCATGGAGA 600
 QY 252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIle 271
 DB 601 ATTCTTGCAGTTCTGGGATTAATGGCGTGPACATTAATTATGATACCGAGTCTCTTTTCATC 660
 QY 272 ProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAlaCysGlyThrSerLeu 291
 DB 661 CCCGAATCCCTCGCTGGCGCAAAATCGGATATGACAGATCAATTTGAAATCTTCAATTA 720

QY 292 GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAla 311
 DB 721 CAAGTCTTCGAGGATTTGAGACTGATATTAACCGTTGAAGTTAATGAATCAAGAGATCT 780
 QY 312 ValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyr 331
 DB 781 GTGGCATCATCTACAAAGCAATACAGTTCGGTTTGTAGATCTCAACCGCAGAGATAC 840
 QY 332 ArgThrProLeuLeuGlyIleGlyLeuValLeuValLeuGlnAsnLeuSerGlyIleAsn 351
 DB 841 TATTTCCCACTTATGGTTGGTATAGGGTTGCTGTACTTCAACAACTGGTGGGAATTAAT 900
 QY 352 GlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeu 371
 DB 901 GGTGTCTTATTCTTCCAGTACAATATTTGAATCTGCAGGAGTTACATCGAGTAACGCA 960
 QY 372 AlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTripleLeu 391
 DB 961 GCACATTTGGGGTCGGGCTATTACAGTAGTAGGACTGCAATATCCACTTGGTTGGTG 1020
 QY 392 AspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeu 411
 DB 1021 GACAAAGCAGGTCTCGCTTCTGTTACTATCTCTCGGTGGGATGACGATTAGCCTT 1080
 QY 412 LeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyr 431
 DB 1081 GTAAATTTGTGAGCTGCTTCTATCTTAAGGAATTTGTCTCTCTGATTCAGACATGTAC 1140
 QY 432 TyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPhe 451
 DB 1141 AGTTGGCTGACATATGTCAGTAGTTGGAGTGTGGCAATGGTTCTTCTTCTCATTTG 1200
 QY 452 GlyMetGlyAlaIleProTrpLeuMetSerGluLeuLeuProValSerIleLysSer 471
 DB 1201 GGAATGGGACCAATACCGTGGCTCATTTATGTCAGATCTCTCTGTGAACATAAAGGTT 1260
 QY 472 LeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThr 491
 DB 1261 TTAGCTGGAAGTATGCAACTCTAGCAATTTGGTTCTTTCTTGGTTGATCACCATGACA 1320
 QY 492 ThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValSer 511
 DB 1321 GCAAAATTTGCTGTAGCTGGAGCAGTGGAGCACTTTCTACTGTATGGATTGGTTGT 1380
 QY 512 AlaPheThrIleValPheValValLeuTrpValProGluThrLysGly***AsnSerArg 531
 DB 1381 GCATTCACAGTGGTGTTCGTGACTCTATGGGTCTCTGAGACCAAGGCAA-AACTCTTGA 1439
 QY 532 GlyAspThrIlePheValSerLeu 539
 DB 1440 AGAATCTCAATCTCTGTTCAGATG 1463

RESULT 9

AF462803

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

REFERENCE

AUTHORS

Chen, R., Chen, H., Kim, C.J., Meyers, M.C., Ban, J., Bower, J.,

Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,

Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,

Lan, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,

Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,

Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K.,

Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A.

and Ecker, J.R.

TITLE Arabidopsis cDNA clones

Unpublished
2 (bases 1 to 1914)
Chen, R., Chen, H., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L.,
Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,
Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,
Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K.,
Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A.,
and Ecker, J.R.
Direct Submission
Submitted (27-DEC-2001) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.

TITLE

JOURNAL

COMMENT

The Salk, Stanford, PSEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Cheuk, R., Chen, H.,
Kim, C.J., Meyers, M.C., Shinn, P., Banh, J., Bowser, L., Dale, J.M.,
Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G., Lam, B.,
Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M., Onodera, C.S.,
Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A., Tang, C.C.,
Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W.,
Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES

Location/Qualifiers
1..1914
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAPL09-48-G03 (R21488)"
/ecotype="Columbia"

ORIGIN

Alignment Scores:
Pred. No.: 8.32e-120 Length: 1914
Score: 1726.50 Matches: 337
Percent Similarity: 83.03% Conservative: 69
Best Local Similarity: 68.92% Mismatches: 75
Query Match: 62.13% Indels: 9
DB: 8 Gaps: 2

US-10-051-909-36 (1-553) x AF462803 (1-1914)

QY 52 GluGluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySer 71
DB 274 GAAGAGGGAAGAATGAT-----CTTGGGAGCCGTTCTTACACACAGGAAGT 321
QY 72 TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal 91
DB 322 TGGTATCGGATCGGGTTCGAGACAACTCTAGTATG-----TTGGAATCGTCTCAAGTT 372
QY 92 LeuArgGluSerHisValSerAlaPheLeuLeuCysThrLeuLeuValAlaLeuGlyProIle 111
DB 373 ATTGAGACAGCTCAATTTCTGCTTACCTGTTGTTGATGTTGCTTCTTGGTCCATATT 432
QY 112 GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu 131
DB 433 CAATTCGGATTCACCTTGTTGGTATTTCCTCACCAACTCAAGCTGCACTTACTTAAGGATCTT 492
QY 132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyMetVal 151
DB 493 GGTTTAACTGATCAGAGTACTCTGTGTTGGGTCTCTATCCAAATGTTGGGTGCTATGGTT 552

QY 152 GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGly-ArgGlyGlySerLeuMetIle 171
DB 553 GGAGCAATTCGAGTGGTTCAGATTCCTGAATACGTTGGACTTGAAGGGTCTCTCATGAT 612
QY 171 exalalalleProAsnIleIleGlyTyrLeuAlaIleSerPheAlaIleAspAlaSerPh 191
DB 613 TGCTGCAATTCCTCAATATTATTGGATGGCTTTCGATATCATTTGCAAAAGATACTTCTTT 672
QY 191 eLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPr 211
DB 673 TCTTTACATGGGAAGATTTAGAAAGGCTTCGAGTTGGGATTATCTCTATACGGTCC 732
QY 211 oValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGl 231
DB 733 CGTATATATAGCTGAGATCGCTCCACAGACCATGAGAGGAGCCTTAGTTCAAGTTAAACA 792
QY 231 nLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrAr 251
DB 793 GCTTTCTGPAACAATTGGGATAATGTCGCGTATTATTACTCGGTCTCTTTTGTTCATGGAG 852
QY 251 gLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPheIle 271
DB 853 AATCTTCGAGTTTGGGAGTATTCCTATGATACATTGTTGATACCGGTCTTTTTCAT 912
QY 271 eProGluSerProA-gTyrLeuAlaIleAspMetAsnLeuThrGluAspCysGluThrSerle 291
DB 913 TCTGAATCTCTCGTGGTGGTTCGCAAGATGGGTTTGACAGATGATTTGCAAACTTCATT 972
QY 291 uGlnValLeuArgGlyPheGluThrAspIleThrGluValAsnAspIleLysArgAl 311
DB 973 GCAAGTTCTTCGTGGATTGAGACTGATATACCGTTGNGGTTTATGAATCAAGAGATC 1032
QY 311 aValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTy 331
DB 1033 TGTGGCATCATCTAGCAAACTTCTGAGTTTGTAGACCTCAACGCGGAGGAGATA 1092
QY 331 rArgThrProLeuLeuGlyIleGlyLeuValLeuGlnAsnLeuSerGlyIleAs 351
DB 1093 CTAATTCACCTGATGGTGGTATAGGCTGCTTGCACTTCAACACTTGGAGGAATCAA 1152
QY 351 nGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLe 371
DB 1153 TGGTGTCTTGTCTATTCGAGTACAATTTTCAATCTCGAGGGGTGACATCAAGTAATGT 1212
QY 371 uAlaThrCysSerSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrLeuLe 391
DB 1213 GCGCACATTTGGAGTTGGCGTTGTTTCAGGTATGGCGACTGGGATAGCCACATGGTTGT 1272
QY 391 uAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysle 411
DB 1273 GGATAAGCAGGTCGTCGGCTTCTACTCATGATCTTCTATCGGAATCAGCATAGCCT 1332
QY 411 uLeuAlaValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTy 431
DB 1333 CGTGATTGTAGCAGTTGCAATTTTACCTTAAGGAATTTGTATCACCTGATTTCAACATGTA 1392
QY 431 rTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPh 451
DB 1393 CAACATTCAGCATGGTTTCGAGTTGGAGTTGGGCTATGGTATTATTCTTCTCTCT 1452
QY 451 eGlyMetGlyAlaIleProTyrLeuMetMetSerGluIleLeuProValSerIleLysSe 471
DB 1453 AGGAATGGGACCAATTCATGGCTGATTTGCTGAGATCTTCCAGTCAATATAAAGGG 1512
QY 471 rLeuGlyGlySerIleAlaThrLeuAlaAsnTyrLeuThrSerPheAlaIleThrMetTh 491
DB 1513 TTTTAGCCGGAAGPATAGCAACTTTGTAACCTGGTTCGTTGATGGTTTATGTTTCTGCT 1572
QY 491 rThrAsnLeuMetLeuThrTyrSerValGlyGlyThrPheLeuSerTyrMetValValSe 511
DB 1573 TGCATAATGCTTTAGCATGGAGCAGTGGGGAACTTTTCACTCTCTACGCTTGTGTTG 1632

QY 511 rAlaPheThrIleValPheValValLeuTTPValProGluThrIleValSerGly***AsnSerAr 531
 Db 1633 TCGATTCACTGGTGTGTTTGTGAGCTTTTGGTCTCCTCAGACTAAAGGAAA-AACGGTTG 1691
 QY 531 gGlyAspThrIlePheValSerLeu 539
 Db 1692 AAGAGATCCAAAGCTTTGTTTCAGATG 1716
 RESULT 10
 AX653119
 LOCUS AX653119 1377 bp DNA linear PAT 22-MAR-2003
 DEFINITION Sequence 2989 from Patent WO03000898.
 ACCESSION AX653119
 KEYWORDS AX653119.1 GI:29155933
 SOURCE Oryza sativa
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 REFERENCE 1
 AUTHORS Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M.,
 Katagiri,F., Quan,S., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.
 TITLE Plant genes involved in defense against pathogens
 JOURNAL Patent: WO 03000898-A 2989 03-JAN-2003;
 Syngenta Participations AG (CH)
 FEATURES
 source 1..1377
 /organism="Oryza sativa"
 /mol_types="unassigned DNA"
 /db_xref="taxon:4530"
 ORIGIN
 Alignment Scores:
 Pred. No.: 9,07e-119 Length: 1377
 Score: 1710.00 Matches: 358
 Percent Similarity: 78.36% Conservative: 44
 Best Local Similarity: 69.79% Mismatches: 39
 Query Match: 61.53% Indels: 73
 DB: 6 Gaps: 9
 US-10-051-909-36 (1-553) x AX653119 (1-1377)
 QY 44 AsnArg---GlyGlyAlaGlyAlaGlyGluSerGlySerAspHisasp-----Gly 60
 Db 4 AACAGGGCGCGCGCGTGGCGTGGCGATGAGAGCGGAAGCGACTACGAGAGCGCGCGC 63
 QY 61 ValLeuArgArgProLeuLeu---AsnThrGlySerThrTyrArgMetSerSerArgGln 79
 Db 64 GGCATGCGGAAGCGCGTGGTATGATGACACAGCGGAGTGTGTACAGGATGGGGTCCGGCGCAG 123
 QY 80 SerSerPhe---AlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSer 98
 Db 124 GGGAGCCTCACCGCGCGCGGAGCTGTCTCATGGCCATCTACCGAGTCCACAGTCTCC 183
 QY 99 AlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGly 118
 Db 184 GCCTTCCTCTCGAGCTCATCGTGGCGTCCGCGCCATCCAGTTCGAGATTCAGGGGGGC 243
 QY 119 PheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPhe 138
 Db 244 TTCTCTCCCGACGAGAGCGCATCATCCGAGACTCGACCTACCTCTCCGAGTTC 303
 QY 139 SerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGln 158
 Db 304 TCGGTGTTCGAGTCTGTCTCAACGTCGGCGCCATGGTTGGGGCGATTCGCAAGTGTCAA 363
 QY 159 MetAlaGluThrIleGlyArgGlyGly----- 167
 Db 364 ATGCCGAGTACATTTGGCGCAAGGKGTAGAAATTTTTTCCTTTTGGGTACACATC 423
 QY 168 SerLeuMetIleAlaAlaIleProAsnIleIleGlyThrLeuAlaIleSerPheAlaLys 187

Db 424 TCATTGATGATTGTCGAATTCCTCAAAACATCATTTGGTGGCTTGGCATCTCCTTTGCCAAAA 483
 QY 188 AspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleSer 207
 Db 484 GACTCATCGTTCTTTATATATGGACGATTGCTCGAGGGGTTTGGTGTGGTGTCTCTCT 543
 QY 208 TyrThrValProValTyrIleAlaGluLeuSerProGlnAsnMetArgGlyAlaLeuGly 227
 Db 544 TATACGGTGGCAGTTTACATAGCAAAATATACCTCAAAACATGAGAGGTGCTCTTGGC 603
 QY 228 SerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPhe 247
 Db 604 TCAGTGAATCAGTATATCTGTAACCGTGTATATATTTGGCATATTTGCTCGGATGTTT 663
 QY 248 IleProTyrArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGly 267
 Db 664 GTTCCTTGGAGGCTTCTTGTCTAATAGGAATCTTGCCTTGCACCTGTGTTCATACCTGGC 723
 QY 268 LeuPhePheIleProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCys 287
 Db 724 CTAATCTTCATTCAGAAATCCCAAGATGGTGGCAAGATGAACATGATCGATGATTTT 783
 QY 288 GluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrGluValAsnAsp 307
 Db 784 GAGACTTCTTTACAAGTTCTCGAGGGATTGAGACTGACATCAGCGCGAAGTGAATGAT 843
 QY 308 IleLysArgAlaValAlaSerSerLysArgThrThrIleSer---PheGlnGluLeu 326
 Db 844 ATAAAG-----GTAGCATCTTCAAGCAGCAGCATATTTTAGCCGATTTGGGTTTCGTG 897
 QY 327 AsnGlnLysLysTyrArgThrProLeuLeuGlyIleGlyLeuValLeuGlnAsn 346
 Db 898 ACACGCAAA----- 906
 QY 347 LeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyVal 366
 Db 907 -----ATTTCTTCAGGTCTC 921
 QY 367 ThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyVal 386
 Db 922 ACAACAGTGTGGCTACATGTGCATGTGCTGCTATCCAG----- 963
 QY 387 ThrThrTrpLeuLeuAspArgAlaGlyArgIleLeuLeuIleIleSerThrSerGly 406
 Db 964 -----ATCTCTTCTGCTGG 978
 QY 407 MetThrLeuCysLeuLeuAlaValSerValValPhePheLeuLysAspAsnIleSerGln 426
 Db 979 ATGACTCTTAAGCTTCTTGGAGTTGCTGTGTATTTTCTTCAAGGATAGCATTTCAAA 1038
 QY 427 AspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheVal 446
 Db 1039 GATTCTCACATGTACTACACCTTAAGTATGATCTCTTGGTGTGCTGTGTGGCTTTGTA 1098
 QY 447 IleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuPro 466
 Db 1099 ATCGCTTCTCTTGGTATGGGTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1158
 QY 467 ValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPhe 486
 Db 1159 GTTAGTATCAAGATCTCGCAGGAAGCTTTCGACGCTCGCCAACTGGCTTACATCCTTT 1218
 QY 487 AlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSer 506
 Db 1219 GGAATAACAATGACAGCAAACTTGTATGCTTGGAGTGTCTCGAGGAGCTTTGTGTCTC 1278
 QY 507 TyrMetValValSerAlaPheThrIleValPheValValLeuTrpValProGluThrLys 526
 Db 1279 TACATGTCGTGAGTGTCTTACCCCTGTGTCTGTATCTTGGTGGTGGCAGACAAAA 1338
 QY 527 Gly***AsnSerArgGlyAspThrIlePheValSerLeu 539
 Db 1339 GGAAG-AACTCTCGAAGAGATACAATGGTCTCTCCGCTG 1376

RESULT 11	BT014234	1825 bp	mRNA	linear	PLN 11-MAY-2004					
LOCUS	Lycopersicon esculentum	clone 133433R,	mRNA	sequence.						
DEFINITION										
ACCESSION	BT014234	1	GI:47105649							
VERSION	BT014234.1	GI:47105649								
KEYWORDS	FLI CDNA.									
SOURCE	Lycopersicon esculentum (tomato)									
ORGANISM	Lycopersicon esculentum									
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;									
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;									
	asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.									
REFERENCE	1. (bases 1 to 1825)									
AUTHORS	Kirkness, E.F., Wang, W. and Vazeille, A.									
TITLE	Direct Submission									
JOURNAL	Submitted (11-MAY-2004) The Institute for Genomic Research, 9712									
	Medical Center Drive, Rockville, MD 20850, USA									
FEATURES	Location/Qualifiers									
source	1..1825									
	/organism="Lycopersicon esculentum"									
	/mol_type="mRNA"									
	/db_xref="taxon:4081"									
	/clone="133433R"									
	/tissue types="fruit"									
	/note="TZHM58"									
ORIGIN										
Alignment Scores:										
Pred. No.:	2,52e-117	Length:	1825							
Score:	1693.00	Matches:	328							
Percent Similarity:	82.85%	Conservative:	73							
Best Local Similarity:	57.77%	Mismatches:	79							
Query Match:	60.92%	Indels:	5							
DB:	8	Gaps:	2							
US-10-051-909-36 (1-553) x BT014234 (1-1825)										
Qy	57	AspHisaspGlyValLeuArgProLeuLeuAsnThrGlySerTptVrArgMetSer	76							
Db	122	AACAAATGAAGGGATCTGAAGAAGCCATTTTACATCTAGGAGTTGGTATCGTATGGG	181							
Qy	77	SerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaValLeuArgGluSerHis	96							
Db	182	TCAAGACAGTCTAGTATG-----ATGGGTCTCTCTCAAGCTATTTCGTGATGTTCT	232							
Qy	97	ValSerAlaPheLeuCysThrLeuValAlaLeuGlyProIleGlnPheGlyPheThr	116							
Db	233	ATTTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT	292							
Qy	117	SerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSer	136							
Db	293	TGTGCTACACTTCAACCAACGCAATTCGAATTCACCAAGATCTGAAGCTCACTGTTCT	352							
Qy	137	GluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSer	156							
Db	353	GAGTCTCTCTTGTCTCTCTTATTCCTCAATGTGGTGTCTGTGTGGTCAATATCTAGT	412							
Qy	157	GlyGlnMetAlaGluTyrIleGlyArgGlySerLeuMetIleAlaAlaIleProAsn	176							
Db	413	GTCAAATTCGAGTACATTTGGCGAAGAGGCTCTTAAATGATAGTCTATACCAAT	472							
Qy	177	IleIleGlyTrpLeuAlaIleSerPheAlaIleSerPheAlaIleSerPheLeuTyrMetGlyArg	196							
Db	473	ATCATTGGTTGGCTTGGCTTCAATTTGCCAAGATCTCTTTCTATACATGGGAAGA	532							
Qy	197	LeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGlu	216							
Db	533	TTGTTGGAGGGCTTGGTGGGAATAATCTTTATGTGGTTCCTGTATATATGCTGAG	592							
Qy	217	IleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPhe	236							
Db	593	ATTGACCTCAGAAATGAGAGGGGCTCTAGGCTCTGTTAACCAAGCTCTCTGTTACAATT	652							

RESULT 12
AX653120
LOCUS
DEFINITION
Sequence 2990 from Patent WO03000898.
ACCESSION
AX653120
VERSION
AX653120.1 GI:29155934
KEYWORDS
Oryza sativa
SOURCE

AX653120 1428 bp DNA linear PAT 22-MAR-2003
Sequence 2990 from Patent WO03000898.
AX653120
GI:29155934
Oryza sativa

237 GlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTrpArgLeuLeuAlaValIle 256
653 GGGATCATGTTTCTTATTTACTGGGACTTTTCTTAATTTGGAGGCTGCTTGTCTTCT 712
257 GlyAlaLeuProCysThrMetLeuIleProGlyLeuPheIleProGluSerProArg 276
713 GGAGCATTTACCTTGCCTAGCACTGATACCTGGACTATTTTTCATCCCGAATCTCTCGG 772
277 TrpLeuAlaIleMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGly 296
773 TGGTTGGCCAAATGGGTTTAAACAGAGATTTTGAACCTTCTGTTGCAAGTTCTTCGAGC 832
297 PheGluThrAspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSer--- 315
833 TTTGAGGCTGATATTACCAATGAAGTAAATGAAGTAAAGAGTCTGTAGCATCATCAACA 892
316 SerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeu 335
893 AACCGAAGACAGCAATGCGTTTTCGGATCTCAACATGAAGATATATGTTTCTCTC 952
336 LeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPhe 355
953 ATGGTAGGAATTTGGACTTACTTATCTCTGCAACAGTTAGGAGGAATCAACGGTGTATATTC 1012
356 TyrAlaSerSerIlePheLysAlaIleGlyValThrAsnSerAspLeuAlaThrCysSer 375
1013 TATTCAGACATTTCTCTATCTCGGCTGATTTCTCAAGTAAATGCTGCAACTTTTGT 1072
376 LeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGly 395
1073 GTCGGTGTCTATACAGTTGTGCTCAACTGCTGAGTTCTTCTACATGCTGCTGAGCAACTGGA 1132
396 ArgArgIleLeuLeuIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSer 415
1133 CGTAGAATTTTATTAATTTCTCTCGGCTGATGCTGTTAGTCTTCTTATTGTTCTCC 1192
416 ValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThr 435
1193 GTTGCCTCTCTCTAAAGGGTATGATCTGAGGATTTACCTCTATGCTGCTGGA 1252
436 MetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAla 455
1253 ATAGTCTCGTAGTGGGTGGTGTTCATGATTTATTTTCTCTCGCTCGAATGGACCA 1312
456 IleProTrpLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySer 475
1313 ATTCATGCTGTTATATGCTGAGATTTCTCAATCAAGATCAAGGTCCTCGCTGGAAGT 1372
476 IleAlaThrIleAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMet 495
1373 ACGGCAACATTTACCACTGCTGTTTCTCGTGGTAAATACAGTACTGCACCGTTGCT 1432
496 LeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIle 515
1433 ATGGCTTGGAGCAGTGGAGAACCTTCACTATTATACATGTTATGTTGCTTTCACCGTG 1492
516 ValPheValValLeuTrpValProGluThrLysGly***AsnSerArgGlyAspThrIle 535
1493 GCGTTCGTGGCAATTTGGGTACCGGAGACGAAAGGAG-AACTCTAGAGGAATTCAGTT 1551
536 PheValSerLeu 539
1552 TTCTTCAGATG 1563

JOURNAL Submitted (23-APR-2004) Institute of Botany, Academia Sinica, 128, Section 2, Academia Road, Nankang, Taipei 11529, Taiwan
COMMENT On Apr 23, 2004 this sequence version replaced gi:208961177. The nucleotide sequence of this BAC clone was generated by combining Monsanto and ASPGC-taiwan sequencing data.

FEATURES
source 1. .204649
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="5"
/clone="OJ1781_H11"

ORIGIN
Alignment Scores:
Pred. No.: 1.58e-96 Length: 204649
Score: 1453.50 Matches: 426
Percent Similarity: 37.12% Conservative: 45
Best Local Similarity: 33.57% Mismatches: 54
Query Match: 52.30% Indels: 747
DB: 8 Gaps: 19

US-10-051-909-36 (1-553) x AC120986 (1-204649)
QY 3 SerSerSerPheArgProAlaGlyLysLysLysLysLysAsnGlnGlyLeuArg 22
Db 118274 TCCTCTCATCCATTCTCGGTGAGCGCGCGC----- 118309
QY 23 ArgGluAlaValProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlyGly 42
Db 118310 -----GCGGAGCGCGAGCGAGCGAGCGCGCGCGCGCGCGCG 118357
QY 43 SerAsnArgGlyGlyAlaGlyAlaGlyGluGluSerGlySerAspHisasp----- 59
Db 118358 AGACGGCGGGAGGTGGTGGTGGCGGACGAGCGCGCGCGCGCGCGCGCGCG 118417
QY 60 ---GlyValLeuArgArgProLeuLeuAsnThrGlySerTrpTyrArgMetSerSerArg 78
Db 118418 GCGCGCGGTGGCGAGCGCGTCTGAACACGCGGAGCTGGTACAGGATGGGATCGCG 118477
QY 79 GluSerSerPheAlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSer 98
Db 118478 TCCAGCTCGCGCGC-----TCTCCATGCGCGCCATCGGGAGTCCACGCTCTCC 118528
QY 99 AlaPheLeuGlyThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGly 118
Db 118529 GCCTTCTCTGACGCTCATCGTGGCGCTCGCGCCCATCCCAATTCCGATTCCACGCGCG 118588
QY 119 PheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGlu--- 137
Db 118589 TTCTCTCACCCACCGAGCGCCATCATCGCGACCTCAAGCTCTCCATCTCCGAGGTC 118648
QY 137 ----- 137
Db 118649 CCGTGCCCGTCCAGCTGCGATCTTGGCTGCTTTTGGCTTCTCTTATTGGCTT 118708
QY 138 -----PheSerAlaPheGlySerLeuSerAs 146
Db 118709 TCTTGTAGTGTGAAGTTTATCTTTCTGATGCACTTCTCGCTTTCGGTTTCGCTGTCAA 118768
QY 146 nValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgGly 166
Db 118769 CGTGGCGCCATGGTGGAGCGATCGGCAGTGGGCGAGATGGCGAGTACATGGCCCGGAA 118828
QY 166 sGly----- 167
Db 118829 AGGGTGAGAAATTGTGCTTCATTTGTCCTTATCTGTCTTCTCCACTAGAAGGCCATCCT 118888
QY 167 ----- 167
Db 118889 GTATCTCGACGACGAAACAATTGATACCTTTGTAGTTGACAGGTCATACGTCTACTGTG 118948

QY 167 ----- 167
Db 118949 TATGATTTCCGGTAACACCAATAATTTGCCATTAAAGAAAGTGAGTGGGATTAGAAATGC 119008
QY 167 ----- 167
Db 119009 TGGTCATTTCTGTGAAGTGTAGCTTGTGAAATATGAAACCCAAATTTTCAGTTCAGT 119068
QY 167 ----- 167
Db 119069 TAGCTATATATTGGGGTTTGACCTTAATTTCAAAGTTGACGAGTCTTTAGTTGTGATT 119128
QY 167 ----- 167
Db 119129 GCCATTGATCTGCAAAAGAGGAGTTAGTATGACAGAAATGTGTAGTGAAGTTTGTGT 119188
QY 167 ----- 167
Db 119189 GACATTTCTTTGGTTCATATCTTCAATCACTTAATCAGATCAGTAGTTAACTACAGTG 119248
QY 168 -----SerLeuMetIleAl 172
Db 119249 AAGCTCTATTATTTTACTCTAATGCTCTGTGATGAATCTCCAGTCGTTGATAATTGC 119308
QY 172 aAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLys----- 187
Db 119309 GCGGTTTCTTAACATCATTTGGTTGGCTTGCCATCTCCTTTGCAAAAGTACGTCACATA 119368
QY 187 ----- 187
Db 119369 TTCTCAAGCTATTTTCTCTGTTGCTCTTTTCACTAGTCTTCTCTCCGCACTT 119428
QY 188 -----AspAlaSerPheLeuTyrMetGlyArgLeuLeuGlyPheGlyValG1 204
Db 119429 TGTTTTTCAGGACGGCTCATTTTATACATGGAGCGCTTCTTGAAGTTTGGTGGTGG 119488
QY 204 yIleIleSerTyrThr----- 209
Db 119489 TGTATATCATATACGCTGTGATTACCACTTCATTTTCATTTTATTGCTGCTAGA 119548
QY 209 ----- 209
Db 119549 ATGTTCTGTGGCTTCACTGCTTGTAGTACTTATCTGACTGAGACTGAGCTGTAGAAT 119608
QY 209 ----- 209
Db 119609 GAGGACAGCATCTTTTGTCTAATGAGCTTTTTCATTTTGTGCGCAGAAGACTTTG 119668
QY 210 -----Va 210
Db 119669 GATCAGTTAATAGTTAGTGCATGTATAAACCCCTTAATGTGTGAATTATATCTACCA 119728
QY 210 lProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAs 230
Db 119729 GCGAGTATACATAGCAGAGATATCTCATCAGAACACAAAGAGGAGCACTTGGCTCCGTGNA 119788
QY 230 nGln----- 231
Db 119789 CCAGGTGTGCTTATCGTCAGAGCATCTTATCTTATGCGAGGAAATTTATGTTGTTT 119848
QY 232 -----LeuSerValThrPheGlyIlePheL 240
Db 119849 ATGTTGGACTTACCATGCACTTTTTCATATGATGTGTGTCCTTACCATTGTTCTCT 119908
QY 240 euAlaTyrLeuLeuGlyMetPheIleProTrpArgLeuLeuAlaValIle----- 256
Db 119909 TGGCTTATTTGTAGCATGTTTGTTCCTTGGAGGCTGCTTCCAGTGATAGGTACAAATC 119968
QY 256 ----- 256
Db 119969 AATACCTGATGCTGAATGTTTCTTATTTAGGATATGTATAATAATAATGAATTTTT 120028
QY 256 ----- 256

Db 120029 AACTTTGGATATTTGCTAGTACCTTTTGGACTTATTGCAATTTGTATGTGGGTTAAATGT 120088
QY 257 -----GlyAlaLeuProC 261
Db 120089 CTGGTACCTCTGATCTGATTCAAATATCTGGGTGGCATGCTGTGTAGGAAGCATCCCAT 120148
QY 261 ysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTriPLeu----- 278
Db 120149 GTACATTTGTAATACCTGCTGCTATTTCTTCATTCGCGAATCCCAAGATGGCTGTGTATGT 120208
QY 278 ----- 278
Db 120209 ATACACCTAAGTTGGATATTGTATCACAAAATCTGAAATCTGTGTAGCAAGCACTTTCTTACA 120268
QY 279 -----AlaIysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuA 295
Db 120269 TCTATATGCGAGCAAAATGAAATGATGATGATTTTGGAGCTTCTCTACAAGTTTGA 120328
QY 295 rgGlyPheGluThrAspIleThrGluValAsnAspIleLys----- 309
Db 120329 GGGGTTTGGAGCTGATATTACTGCAGAGTGAATGATATAAAGGCAAGCAATTTTATTT 120388
QY 309 ----- 309
Db 120389 TTCTTGTAACTGACTACAAAGTTATTTGCGGTTTCGCGTGAATTCATCTTAGTAAGATT 120448
QY 309 ----- 309
Db 120449 TTCCTAACACAACTTGGAGATATGTACTTTTCATTTTTCTGGAGACTAATATCGTT 120508
QY 310 -----ArgAlaValAlaSerSerLysArgThrThrIleSerPhe 323
Db 120509 CAATCCCTGTTATGCGAGAGCAGTAGCATCGGCAACAAAGAACCAACAGTCCGTTT 120568
QY 324 GlnGluLeuAsnGlnLysLysThrArgThrProLeuLeuLeu----- 337
Db 120569 AAAGAGTTGAACCAAGAAATACCGCACTCCCTTACTGCT-AACTTACATACCCCTTTT 120627
QY 337 ----- 337
Db 120628 GTTCAGTTTATACAAAGTCATTAATCATGTTTGTAAAGTGAACCTAGCTTATGTTGA 120687
QY 338 -----GlyIleGlyLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPhe 355
Db 120688 CAGATAGGAACCTGGCCCTTCTGTACTTCAGATCTTAAGTGAATAAATGTTATCTGTTT 120747
QY 356 TyrAlaSerSerIlePheLysAlaIleGly----- 365
Db 120748 TATGCAAGTAGAATCTTCAGAGATGCGAGGTAGTGTGGAATATCTCTTATAGTCCCTTAG 120807
QY 365 ----- 365
Db 120808 TACTACAGTCACGTTGTAGTAATTTTGGCTTGTAACTTCTTAACACCAAGAACTTCAGG 120867
QY 366 val-ThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGln----- 380
Db 120868 GTTTACAAACAGTGACTTGGCCCATGTGCACATGTCACTTAGGCAATTCAGGTAAGTATAGGA 120927
QY 380 ----- 380
Db 120928 ATATCTGTATCTCTGTAACTTATATCAAGTTGCTTACAGAAATCGTTGGATTACA 120987
QY 381 -----ValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgIleLe 399
Db 120988 TTAGGTCTTGGCAGCTGGAGTTACACATGTTTACTGGACAGAGCTGGTCGAGCATGT 121047
QY 399 uLeuIle----- 401
Db 121048 CCTTATGTAAATTCCTCTCTCCCAATGAATCACCTTTTGTAGTTGTCTGTACAAATA 121107
QY 402 -----IleSerThrSerGlyMetThrLeuCysLeuLeuA 413

Db 121108 ACATAGTGACTCACTTAAATTCAGATCTCTACTGCTGGATGACTCTAAGCCTTTCTTG 121167
QY 413 lvalSerValValPhePheLeu----- 420
Db 121168 CAGTATCTGTTGTATTTTTTCTGGAGGTACATTCAGATTTTTCAGATATACCATTTGA 121227
QY 420 ----- 420
Db 121228 TAAAAAAATCATATGTTGCACTCTCTTTGATAGCCTACATGTGCGCATGCACTCTCC 121287
QY 421 -----LysAspA 423
Db 121289 CTTTCTCTCAGTTCACCAATATTTCAATATTACTACAATGATTATTGTTTCAGGGTA 121347
QY 423 snIleSerGlnAspSerAsnSerTyrrIleLeuThrMetIleSerLeuValGlyIleV 443
Db 121348 ATATTTCACATGATTTCTCATTCGTTCTACATCTTAAGTATGATCTCTCTGTTGCTCTTG 121407
QY 443 al----- 443
Db 121408 TGGTATGCTGTCAAGTTAAACCATTTCTTTGAAAAATTCCTCCAAAACTCTCGCGGAAA 121467
QY 444 -----SerPheValIleThrPheSerPheGlyMetGlyAla 455
Db 121468 ATTTACCTCGGTCACTTACTTACAGCTTATATCATCACCCTTTCTTCGCGCATGGTGCC 121527
QY 456 ileProTrpLeuMetMetSerGluile----- 464
Db 121528 ATTCCATGGGTCAATGATGCTGTAGGT-ATAATGCTTTGCCCAATAGCTACACATGATGTT 121586
QY 464 ----- 464
Db 121587 AGTATTTTAGCATTGACACATCCACCATTGTTGTCAGATATCTCAGGCATCAGCTT 121646
QY 465 -----LeuProValSerile 469
Db 121647 AAGTAATCTAAAGATTTTCAAGTAAATCTTTGTTGTTCTTAGATCTCTCCAGTTAGCATC 121706
QY 470 LysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThr 489
Db 121707 AAGAGTCTTGGGGAAGCTTTGCGACACTTGCACACATGCTTACATCTCGGCAATAACA 121766
QY 490 MetThrThrAsnLeuMetLeuThrTrpSerValGly----- 501
Db 121767 ATGACAGCAATTTGTTACTCAGCTGGAGTCTGGAGTCCGTTATTTCTATAAAATCCT 121826
QY 501 ----- 501
Db 121827 CTGTACCAAGATAAATAATCTGAACATGATTACCTGATAAAAAGTTAAATATTTTATTT 121886
QY 502 -----GlyThrPheLeuSerTyrrMetValValSerAlaPheThrIleValPheVa 518
Db 121887 AATTTCACAGGAACATTTTTCCTACATGATGTGAGTGCTTTCACTCTCGGTTCGT 121946
QY 518 lvalLeuTrpValProGluThrLysGly***AsnSerArgGlyAspThrIlePheValSe 538
Db 121947 CATATTTGGGTGCGAGACAAAGGAAG-AACTCTGGAGGAGATACAGTTTTCGTCTCT 122005
QY 538 rLeuSerIleGln 542
Db 122006 GCTAAATTTTGAG 122018
RESULT 14
AC025808/c
LOCUS
DEFINITION Genomic sequence for Arabidopsis thaliana BAC F18014 from
chromosome I, complete sequence.
ACCESSION AC025808
VERSION AC025808.8 GI:7636235
KEYWORDS HTG.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
AUTHORS
Shinn, P., Kim, C., Altafi, H., Bei, Q., Chinn, C., Chinn, P., Choi, E., Conn, L., Conway, A., Gonzales, A., Hansen, N., Howing, B., Koo, T., Lam, B., Lee, J., Lenz, C., Li, J., Liu, A., Liu, K., Liu, S., Mukharsky, N., Nguyen, M., Palm, C., Pham, P., Sakano, H., Schwartz, J., Southwick, A., Thaveri, A., Toriumi, M., Vaysberg, M., Yu, G., Federspiel, N.A., Theologis, A., and Ecker, J.R.

TITLE
Genomic sequence for Arabidopsis thaliana BAC F18014 from chromosome 1

JOURNAL
Unpublished

REFERENCE
AUTHORS
Ecker, J.R.

TITLE
Direct Submission

JOURNAL
Submitted (15-MAR-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA

REFERENCE
AUTHORS
Ecker, J.R.

TITLE
Direct Submission

JOURNAL
Submitted (22-APR-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA

REFERENCE
AUTHORS
Cheuk, R., Shinn, P., Brooks, S., Buehler, E., Chao, Q., Johnson-Hopson, C., Khan, S., Kim, C., Altafi, H., Bei, B., Chinn, C., Chinn, P., Choi, E., Conn, L., Conway, A., Gonzales, A., Hansen, N., Howing, B., Koo, T., Lam, B., Lee, J., Lenz, C., Li, J., Liu, A., Liu, K., Liu, S., Mukharsky, N., Nguyen, M., Palm, C., Pham, P., Sakano, H., Schwartz, J., Southwick, A., Thaveri, A., Toriumi, M., Vaysberg, M., Yu, G., Davis, R., Federspiel, N., Theologis, A., and Ecker, J.R.

TITLE
Direct Submission

JOURNAL
Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA

REFERENCE
AUTHORS
Chao, Q., Brooks, S., Buehler, E., Johnson-Hopson, C., Khan, S., Kim, C., Shinn, P., Altafi, H., Bei, B., Chinn, C., Chinn, P., Choi, E., Conn, L., Conway, A., Gonzales, A., Hansen, N., Howing, B., Koo, T., Lam, B., Lee, J., Lenz, C., Li, J., Liu, A., Liu, K., Liu, S., Mukharsky, N., Nguyen, M., Palm, C., Pham, P., Sakano, H., Schwartz, J., Southwick, A., Thaveri, A., Toriumi, M., Vaysberg, M., Yu, G., Davis, R., Federspiel, N., Theologis, A., and Ecker, J.R.

TITLE
Direct Submission

JOURNAL
Submitted (11-OCT-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA

COMMENT
On Apr 22, 2000 this sequence version replaced gi:7543634.

FEATURES
source
1. 120977
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="1"
/clones="F18014"
90. 2159
/note="hypothetical protein"
/codon_start=1
/evidence=not experimental
/product="F18014.1"
/protein_id="AAF79419.1"
/db_xref="GI:8778411"

CDS
translation="MIGFEVSPDVFTCSIVVNAVCRSNVDKAMVFAKETESSLGLEL
NVNTYSLINGYAMIGVEGTVRLRLMSGRVSRNVVTYSLIKGVCKGLMEAEH
VFELKSKLVADHMYGLMDGVCPTGQIDRAVVDHNNIEIGVNTNTINSLNG
YKSGQVQAEQIFSRVNDKSLPDHNTYTLVDGCRAGYVDEALKLQDMQKRW
PYVMYVILLKYSRIGAFHDVLSLWQMLKRGVNADEISCTILLKFKIGQFNAM
KLWENVLARGLTDITLTYNMIISGLCKMEKNEAKETLDNVDNFRCKPAVQTYQALSH
GYKVGKLEKAFVYEMERKGIPTTEMNTYITISGAFKYRHLNKVADIVIELRAGSL
TPTVATYCALITGWCNTGMIDKAYATCFEMIEKGITLNNVICSNIANSIFRLDKIDEA
CILLQKIVDFDLPLPGVQSLKEFLEASATTCLTKQKIAESVNSTPKLLVPPNIVN

VAIAGLCAGKLEADARKLFDLSSDRFIPDEVTYTLIHGCAIAGDINKAFTLRDEM
ALKGIIPNIVYINAIKGLCKLGNQDVAQRIHLKLPQKGIPTNATYTLNIDGLVKS
NVAAEMLKEMIEKGLVRGSDXQGVDDPKEVVLDPEVLGSGTVIEMSNELYDVR
RVSEAVV"

complement (4647. 5702)
/note="putative glycosyltransferase-like protein
emb|CA942905.1; similar to EST gb|AI998490.1"
/codon_start=1
/evidence=not experimental
/product="F18014.2"
/protein_id="AAF79456.1"
/db_xref="GI:8778448"

/translation="MSQHLILLILLILLHLPISATTIIQKPEAPOFVNSADCPCLI
DSSDDDVAKPIFCSPRAVHVAVTLDAAVIRGSVAALVLSQHSPPENIVFHFVA
SASADASLRATISSSPYLDFTYVFNVSVSRLISSIISALDCLNVAERSVLADL
LPCVRRVYLDSDLLIIVDDIAKADLGRDSVLRAPEYCNANFTSYFTTFWNSPT
LSLTFADKACFYNTGVNIDLSRWREGAYTRIEEWMAMKRTIIEILGSIPEFLV
FAGLIKPVNHNQHLGDNFRGLDLHPGVSLDLHWSGKGKFWARLDAGRPCPLD
ALWAPYDLIQTFFALDS"

complement (9620. 10300)
/note="similar to Zn finger protein gi|6319911"
/codon_start=1
/evidence=not experimental
/product="F18014.3"
/protein_id="AAF79455.1"
/db_xref="GI:8778447"

/translation="MSDPVSCSGNDNNDSNFECNICLDLAQDPITVLCGHFLFCW
PCLYKWLHLSQKDCVCXAVIEDRLVPLVYGRKSSADPRSIPGLVFNPSGO
RPFATQDPDHPGFAHHGFCGFMGCPAASAEFGNVTLSAAGGLISLENLHPH
GVFPAALGTGAASGCFPHGFSNPPFHGHSHVHSQRTGRCGQQDHLHLLILVFFV
VVFSLFUS"

join (12523. 12577, 12686. 13374)
/note="hypothetical protein"
/codon_start=1
/evidence=not experimental
/product="F18014.4"
/protein_id="AAF79420.1"
/db_xref="GI:8778412"

/translation="MAIFSTSHLLFISFIATCTISVSGTFTTLNHCOSTIWPGILT
ANGALGCGFALASGVSFTVPSGSRFWARTYCNFDSGSGCKGTGCGSKLKC
AGAGAPPAETAEFTIGSSGKKNAQVDPYDVLVDYGVNOMKI PQGSGSGCKTAGCV
SDVNAICPKELQVTGSPCAKACAFNKPEYCCGTGAYSTPATCPTPTNYSKIFQA
CPSAIYAYDASSTFTCTNANYISFCS"

complement (14020. 14136, 14236. 14304, 14402. 14464,
14557. 14696, 14879. 15005, 16274. 16486)
/note="unknown protein; similar to EST gb|H76493.1"
/codon_start=1
/evidence=not experimental
/product="F18014.5"
/protein_id="AAF79454.1"
/db_xref="GI:8778446"

/translation="MLEAVDSGVYVNGGFPQIQSYGDCSSEELSVLPRTKVVVTG
NNRTSLVGLQGVVVKAVGLGQHWLVTNGIEVKLQNALSVLEPTTGNEDDDLD
PENTORSGDMTSEDTLKPKHKLGRSSSSSHKMSRSLSSQSKSGFTPPENM
KVDLSKLEMPALLANYRHFNLVDAIPNFSKEQLIDIVQHFMSQQMDELQVIVGVQA
ARMKKKAFQSKSRWTDLNCIS"

join (17994. 18489, 18568. 18698, 18814. 18847, 19006. 19099,
19170. 19347, 19441. 19479, 19616. 19715, 19855. 19994)
/note="hypothetical protein"
/codon_start=1
/evidence=not experimental
/product="F18014.6"
/protein_id="AAF79421.1"
/db_xref="GI:8778413"

/translation="MAKTDKLAQFLDSGIYSEDEFNWFLLDTVTRTNSTYRPFKVS
AYSFYFFSKQLNHSSENPKRKKOKNSFHLPSVGEQASLRHQEALFLSKAH
BSFLKEILLSTKGLSDNDSSLLNKCCDEVSFELGGVWQAPFETITLNL
HCDNGBESCNORFQVFNVLVNEIGEEVAEPNRRYIMPRNSCFYMSDLHFRNL
VPAKSEBNYLLIVDPKENASAKHKSKEVLYMYPTLPCNOYFLSLPIKOLAHAGAL
VALWYNREKLSLSEVEKELPFWAGIKYVATWYWLKVKPDGLICLDLLAGSEKSDP
KLLDNQKQIMSLPGDFSRKPPFLLKLLKPGSGQARCELFAREMAAGTWSGNEPL
HPQDSRYFLKV"

join (22037. 22300, 22618. 23361)

CDS

/note="unknown protein; similar to ESTs gb|AI992723.1, gb|T41777.1, dbj|AV440980.1, gb|AI997341.1, and dbj|AV440785.1"
 /codon_start=1
 /evidence=not experimental
 /product="F18014.7"
 /protein_id="AAF79422.1"
 /db_xref="GI:8778414"
 /translation="MTSDGATSAASAAAAMATRKPSWRRENNRRERRRAVAA
 KIVTGLRAQGNVLPKCDNNEVLKALCEAGHVEEDGTVEKHKPLPLGDMAGSS
 RATPYSHNSQPSLSSFDSPILSYQVSPSSSPSRVGDPHNISTIPFLRNGIP
 SSUPPUSINSAVTPVPSPTSRNPKPIPTWESFTQSNMAAKQSMSTLNYFFVAV
 SAPASPHQHPAPATIECDSDSDTVDGSHWISFQKFAQQPFASMSVPTFN
 LVKPAPOOLSPNTAAIQEIGSESEFFKNSQVKEGRIHDVAMEDELETLGNGKAK
 S"
 complement(join(23868..24983,25237..25407))
 /note="unknown protein; similar to EST gb|T13858.1, emb|Z38050.1, and emb|Z38087.1"
 /codon_start=1
 /evidence=not experimental
 /product="F18014.8"
 /protein_id="AAF79453.1"
 /db_xref="GI:8778445"
 /translation="MAGRRDRSQQLRGSRITAILIGIGCVCAVLPPYGFNSSS
 LKASRLKSSNVQGSACSPKRVKMLKSDFTLSEKNELKKQRELTEKRLAEQ
 GSNARQVIALGTQIKAGFPFVKSRLNPTILPDESINPLAKILEBIAVDKRVIV
 ALANAVKAMLEVCQIASIKRVGITYNLVVALDDYIENLCKENDVAYIKSDPKDVTY
 KTGCGHVAVGLFVRLEFQLQGYGVLLSDVDIVLQNPFSHLVYRDSVESMSGDH
 NHTQAFNDVDFDPAMPWARYAHTWIFVNSGFYLRPTIPISIELLDRAVRLSKAK
 WPDQAVNEELFVPSHPEYTAHASKRVMDVEFMNSKVLFTKVRKNHELKKVKPVI
 VHYVHPDLKRNQAVVEYVNGKQDALDSFDGSE"
 complement(26100..27917)
 /note="unknown protein; similar to ESTs gb|T42717.1, emb|F13886.1, and emb|F13885.1"
 /codon_start=1
 /evidence=not experimental
 /product="F18014.9"
 /protein_id="AAF79452.1"
 /db_xref="GI:8778444"
 /translation="MADPLNGKFFFCIFLLSFTLLFSPLYATESPVIEDVSDVA
 VASEYTRAVLHNEELVKNLDELVANLDAKLSATPLKXNEISVDDIGEEKERG
 RAKAFSVTKSPWSERFQTSKAVLNDAKTNVLPFRDFEGSKSYFAIGDKGRVY
 VFLRNGDVLTEFFTTVDSPVTAIVSYVSKNSVFWTGHQNGAVLLHRIHESNGED
 WNSNSVMEHVGVKDFVDDSDADPVTLELVHVGVRVYLATDLSGLTVLTENRTVYGS
 VIFSSRPLVFLKQRLFLFESGAGSLDLSMKIRTECEGLNHLNLTATVTFDAERSK

CDS

Alignment Scores:

Pred. No.:	3,64e-73	Length:	120977
Score:	1137.50	Matches:	335
Percent Similarity:	40.58%	Conservative:	68
Best Local Similarity:	33.74%	Mismatches:	518
Query Match:	40.93%	Indels:	18
DB:	8	Gaps:	

US-10-051-909-36 (1-553) x AC025808 (1-120977)

QY	52	GluGlySerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySer	71
DB	67805	GAAGAGGGAAGAATGAT-----CITCGCGGACCGTCTTACACACAGGAAGT	67758
QY	72	TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal	91
DB	67757	TGCTATCGGATGGGTCGAGACAATCTAGTATG-----TTGGAATCGTCTCAAGTT	67707
QY	92	LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIle	111
DB	67706	ATTCGAGACAGCTCAATTTCTGTCTTAGCTTGTTTGTATGTTGCTCTCTGCTCTAT	67647
QY	112	GlnPheGlyPheThrSer-----	117
DB	67646	CAATTCGGATTCACTGT--AAGAACACATACAAACATAATCTCAACTTCTTATATTTT	67598
QY	117	-----IleGlyAlaLeuProCysThrMetLeuIleProGlyL	268

DB	67587	TTTATCATGTGGGATTTCTCTGCTTTTGCTCAAATTCAGTAAGAAAGTGTGATTTTGAT	67528
QY	118	-----GlyPheSerSerProThrGlnAspAlaMetValArgAsp	130
DB	67527	TTATCTTTTCTTTTTCAGTGTGGTTATCTCTCAACAACTCAAGCTGCATTACTAAGAT	67468
QY	131	LeuAsnLeuSerIleSerGlu-----	137
DB	67467	CTTGCTTAACCTGTATCAGAGGTATGAAGTCCCATCTTTATTTAAGAACTGTTCTGTTT	67408
QY	137	-----	137
DB	67407	TTTTTTTTTTTTTTGTAAGAACTCTTGTTCTGTAATTTTCTGCACTTTTAGGTTCTTGT	67348
QY	138	-----PheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAla	153
DB	67347	TGATGTTTTAGTACTCTGCTGTTTGGGTCTCTATCCAAATGTGGGTGCTATGTTGGAGCA	67288
QY	154	IleAlaSerGlyGlnMetAlaGluTyrIleGlyArg-LysGly-----	167
DB	67287	ATTGCGAGTGTGAGATTGCTGAATGCTTGACGGAAGGGGTATAATAGTTTGTGTTTC	67228
QY	167	-----	167
DB	67227	TGATTTCTAAACANGTTCTCATCTGTGACAATTTGTCATGTAACATCTTCTTCTGTTT	67168
QY	168	-----SerLeuMetIleAlaIlePr	175
DB	67167	GTATCTTATTTACTGTGTTTTCTTCTTCTTGTGACAGTCTCTGATGATTGCTGCAATTC	67108
QY	175	AsnIleIleGlyTrpLeuAlaIleSerPheAla-Lys-----	187
DB	67107	CAATATTATTGGATGGCTTTTTCGATATCATTTGCAAAAGTCAAGTGTGTTGTTGTTCTCAA	67048
QY	187	-----	187
DB	67047	GTTTCTGATATTTCTGGCTTATCTCTATACCTACTAGTTGTTGTTGTTCTTGGTTTTTG	66988
QY	188	-----AspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleI	206
DB	66987	TTTAGGATATCTTCTTTTTCATATGGAAGATTGTTAGAAGGCTTCGGAGTTGGGATTA	66928
QY	206	IleSerTyrThrVal-----	210
DB	66927	TCTCTTATACGGT--AATATTAGATTCAACTATTCTTTGATTTCAGCAATGTTTGTGG	66869
QY	211	-----ProValTyrIleAlaG	216
DB	66868	AGTGATCAATATAATTACACACCTTTTGAATTCGATTTTAGGTTCCGATATATATAGCTG	66809
QY	216	IuIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeu-----	232
DB	66808	AGATCGCTCCACAGACCATGAGAGGAGCCTTAGGTTCAAGTTAACAGGT--TTGTGATGTT	66750
QY	232	-----	232
DB	66749	TAGTTAGCTTCACTATTTTTATGTTTTCATTTTATGAGCTGATTCAATGTTAAACATAAA	66690
QY	233	-----SerValThrPheGlyIlePheLeuAlaTyrIleuL	244
DB	66689	CAAPAAATGTTTGTGTTTATAGCTTTCTCTATACAATTTGGGTAATGCTGCGGATTTAC	66630
QY	244	euGlyMetPheIleProTyrArgLeuLeuAlaVal-----	255
DB	66629	TCGCTCTCTTGTGTTTCCATGAGGAATCTTCGAGTTTGGGTAACTCACAAACATTTTGA	66570
QY	255	-----	255
DB	66569	CTTGACTTGGGAGAACTATACCTGTTTATTTTATTTTATGTCGATACAGTCTAATAC	66510
QY	256	-----IleGlyAlaLeuProCysThrMetLeuIleProGlyL	268

Db	66509	TTCTCTTTGTTGTTCTTTATGTTAGGAGTATTGCCATGTACATATGTTGTATACCGGGTC	66450
Qy	268	euphepelleProGluSerProArgTrpLeu	278
Db	66449	TTTTTTTCATTCCTGAATCCCTCGGTGGTGGTATGCTCTACAAACCGTATTGCTTTATA	66390
Qy	279	-----Ala	279
Db	66389	TTGTTGAGTTGTTGTCATTATTGGTTATCTCTAAATGTTTGGTGTGTTGTAATTGTAGGCA	66330
Qy	280	LysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThr	299
Db	66329	AAGATGGGTTTGACAGATGATTTCGAAACTTCATTCGAAATGTTCTTCGTGGAATTTGAGACT	66270
Qy	300	AspIleThrThrGluValAsnAspIleLys	309
Db	66269	GATATTACCGTTGAGGTTATGAAATCAAGTAGTACTAACTGAAGAAATGCTTTGAT	66210
Qy	310	-----ArgAlaValAlaSe	314
Db	66209	GTGATTTCTTGTCTTATTTAAATTTTTCATTTTATGTCACCTCTGCAGAGATCTGTGCATC	66150
Qy	314	rSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysrArgThrPr	334
Db	66149	ATCTAGCAAAACGTTCTGCAGTTTCGGTTGTAGACCTCAAGCGCAGGAGATACATTATCCG	66090
Qy	334	oLeuLeu	336
Db	66089	ACTGATGGTGCATTTCTCTTACTCTCTTGAACATAATTTTGTCAACACATGAGAGAAC	66030
Qy	337	-----LeuGlyIleGlyLeuLeuValLeuGlnAsnL	347
Db	66029	TCACTCTTTTCGTAAATCGTGTACAGGTTGGTATAGGGCTGTGCACCTTCAACAAC	65970
Qy	347	euSerGlyIleAsnGlyValLeuPheThrAlaSerSerIlePheLysAlaAla	364
Db	65969	TTGGAGGAAACAATGGTGTCTTCTTATTCGAGTACAAATTTTGAATCTGC-AGGTAGA	65911
Qy	364	-----	364
Db	65910	GTCAATATGTTTTTTTGTGATTTGGAATGTCCTAAGATAGAATAAACAGATTTCTAACT	65851
Qy	365	-----GlyValThrAsnSerAspLeuAlaThrC	374
Db	65950	ATGCTTCTTTGTTGCAATATCATTCCTCAAGGGTGACATCAAGTAATGTGGCAGAT	65791
Qy	374	ysSerLeuGlyAlaIleGlnValLeu	382
Db	65790	TTGGAGTTGGCTGTGTTTCAGGTATA-TCATTTCAATCTAGGAGACATTAGATTCGATTTT	65732
Qy	383	-----AlaThrGlyValThrThrTrpLeuLeuA	392
Db	65731	TCTAACCGTTGTTTCATTTATACAGGTAGTGGGACTGGGATAGCAGATGTTGGTGG	65672
Qy	392	spArgAlaGlyArgArgIleLeuLeuIle	401
Db	65671	ATAAGCAGTCTCGGCTTCTACTCATGTGTGAAGATATTTGAGATTCATCTGTATATC	65612
Qy	401	-----	401
Db	65611	TTCAAAAGTCTGAAGAATTTATGTATGCATAACAATAATATTAGCATTTTCTCTGAAATG	65552
Qy	402	---IleSerThrSerGlyMetThrLeuCysSerLeuAlaValSerValValPhePheLeu	420
Db	65551	CAGATCTCTTCTATCGGAATGACGAAAGGCTCGTGNATTGTAGCAGTTGCAATTTACCTT	65492
Qy	421	Lys	421
Db	65491	AAGTAAACATTTTAACTTTTGACATCTCTACTTGTACATATGTAAGAGTGTCACTCTAAT	65432
Qy	422	-----AspAsnIleSerGlnAspSerAsnSer	430
Db	65431	ATGAGATATTGAGGTTTCTTAACTTTTGTAGGAATTTGTATCATCCTGATTCACACATG	65372

Qy	431	TyTrpIleLeuThrMetIleSerLeuValGlyIleVal	443
Db	65371	TACAACATTTCTAGCATGGTTTCGATGTTGGAGTTGGTAAGTCGTTGTTTATACA	65312
Qy	443	-----	443
Db	65311	TTCACACAACAAAGATACTTTCTTATTTACATATTTCCCGGTATGTGTACTAAATCTGG	65252
Qy	444	-----SerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrp	458
Db	65251	CTGAGTTATTACAGGCTATGTTTCTTCTCTCTAGGAATGGGACCAATTCATCG	65192
Qy	459	LeuMetMetSerGlu	463
Db	65191	CTGATTATGTCGAGGTACTAAATATTAAACATGGTCTTCTCTAGTGTAAATACAA	65132
Qy	464	-----IleLeuProValSerIleLysSe	471
Db	65131	CTGGAAATTTTAATACTGATTAAATGTTGGTAAACAGATCTGCCAGTGAATATAAGGG	65072
Qy	471	rLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetTh	491
Db	65071	TTTACCGGAAGTATAGCAACTTTGTTAACTGGTTCGTGTCATGGTTAGTGACATGAC	65012
Qy	491	rThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTrpMetValValSe	511
Db	65011	TGCAATATATGCTGTAGCATGAGCAGTGGGGAACCTTCACTCTCTACGTTTGGTTG	64952
Qy	511	rAlaPheThrIleValPheValValLeuTrpValProGluThrLysGly***AsnSerAr	531
Db	64951	TGGATTCAGTGGTTTTTGTGAGCTTTTGGTCTCTGAGACTAAAGGAAA-AACGCTTG	64893
Qy	531	gGlyAspThrIlePheValSerLeu 539	
Db	64892	AAGAGATCCAAGCTTTGTTCAGATG 64868	
RESULT 15			
AC025814			
LOCUS	68041 bp	DNA	linear
DEFINITION	Arabidopsis thaliana chromosome 1 BAC F22H5 genomic sequence,		PLN 22-JAN-2001
ACCESSION	AC025814		
VERSION	AC025814.7	GI:12331602	
KEYWORDS	HTG.		
SOURCE	Arabidopsis thaliana (chale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1 (bases 1 to 68041)		
AUTHORS	Lin,X., Kaul,S., Town,C.D., Benito,M.-I., Creasy,T.H., Haas,B.J., Wu,D., Maiti,R., Ronning,C.M., Koo,H., Fujii,C.Y., Utterback,T.R., Barnstead,M.E., Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.		
TITLE	Arabidopsis thaliana chromosome 1 BAC F22H5 genomic sequence		
REFERENCE	2 (bases 1 to 68041)		
AUTHORS	Lin,X. and Kaul,S.		
JOURNAL	Direct Submission		
TITLE	Medical Center Dr, Rockville, MD 20850, USA, xlin@igr.org		
REFERENCE	3 (bases 1 to 68041)		
AUTHORS	Town,C.D. and Kaul,S.		
JOURNAL	Submitted (12-SEP-2000) The Institute for Genomic Research, 9712		
TITLE	Medical Center Dr, Rockville, MD 20850, USA, cdtown@igr.org		
REFERENCE	4 (bases 1 to 68041)		
AUTHORS	Town,C.D. and Kaul,S.		
JOURNAL	Submitted (05-OCT-2000) The Institute for Genomic Research, 9712		
TITLE	Medical Center Dr, Rockville, MD 20850, USA, cdtown@igr.org		
REFERENCE	5 (bases 1 to 68041)		
AUTHORS	Town,C.D. and Kaul,S.		

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (12-OCT-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
6 (bases 1 to 68041)
Town, C.D. PhD.
Direct Submission
Submitted (22-JAN-2001)
On Jan 22, 2001 this sequence version replaced gi:12280794.
Address all correspondence to: at@tigr.org

BAC clone F22H5 is from Arabidopsis thaliana chromosome 1
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.

Genes were identified by a combination of several methods: Gene
prediction programs including Genscan+ (Chris Burge,
<http://CCR-081.mit.edu/GENSCAN.html>), GeneMarkHM (Mark Borodovsky,
<http://genemark.biology.gatech.edu/GeneMark/>), GlimmerA (a variant
of GlimmerM, see Mihaseia Pextea,
<http://www.tigr.org/softlab/glimmerm.htm>/glimmerm.html, and
GeneSplicer (Mihaseia Pextea and Steven Salzberg, contact
mpextea@tigr.org), searches of the complete sequence against a
peptide database and the plant EST database at TIGR
(<http://www.tigr.org/tdb/tgi.shtml>). Annotated genes are named to
indicate the level of evidence for their annotation. Genes with
similarity to other proteins are named after the database hits.
Genes without significant peptide similarity but with EST
similarity are named as unknown proteins. Genes without protein
or EST similarity, that are predicted by more than two gene
prediction programs over most of their length are annotated as
hypothetical proteins. Genes encoding tRNAs are predicted by
tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>).
Simple repeats are identified by repeatmasker (Arian Smit,
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>).

FEATURES

source
1. 68041
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="1"
/clone="F22H5"
complement (418..5404)
/gene="F22H5.15"
complement (join(<418..594,684..773,873..938,1034..1095,
1179..1227,1319..4820,5040..5404))
/gene="F22H5.15"
complement (join(<418..594,684..773,873..938,1034..1095,
1179..1227,1319..4820,5040..5404))
/gene="F22H5.15"
/codon_start=1
/product="unknown protein, 3' partial; 5120-134"
/protein_id="AAG12690.1"
/db_xref="GI:10092277"
/translation="MEVEKSPATTPSRKISNNRSHLSFSANAVYDGVFFSPVNSKS
PLVDYGEIFRGSFSPSSIPFLDVPALVGVKVDVRSKLDYSSVGGSLGACDFAVT
PKVILIKSKTSINEDKRRKNGSDVPLCNCKKSPKRWKRKSDISYHQTVP
RNNGATHLPQVATPPIPTQVDTNLSLHKIESKTPIPAVERKKUPCNEGREVK
ASRQKSTVEDFADGCTDSTCTVNGEYDVRKPPSFQCTLNHGAS
ERLGLNSGSPSEYEDADSPSPYFADNTDENSAAESSAALKAEAEQIRMI
AKOMKKKSGFRSKLSCDDSKENKNTKVEGITEEDNNSQLGEMVKPSEQ
SFNEDQAKARKLWGPGLKSTDPKPEELEEDQITILEEQAARRGRKHWELP
GGTPKSMVSKQEPENLAPKPEPTQKQVOPITENPYTFQGLSKLKVCEAFTG
SKVSKDEKQFTIKENSTVQVQDESSQMLAGIPVIEYLRVEETPQTESKS
EMNIEKSESTICAFTRSSQNKRETKGQVSAKCPDGSGKQVQFQNGDHTCNVLD
QEGEIVSEPOMLVGPDSKTYREVEETPTSLNKQSDSDVGAMVFNVRNISE
PGNIDVQAVHVRPVRVWKTSEDVYNNIKAPKNRFPWQLEAGNETTCLVSEEG
GVRIHASEIEISTSQSDSLQENWTVLKMFRMOTADTKGEDTCTYCLVSEEG
HLDIHOKAQKQVQVETVRTNYEAYAHTRNEDESAQETCYCKEDGVEYQGRKTSIL
VRLLIGEELEMSLEDEYQAEAGVQGLSELNEINHADSHAEMLLEYDRSTDS
NNSRFPDQEQAESTMIDGSDTTSRSSFEMRQGDYSIEVGLQHRHQPFPEKA
TSVGNTEHIEEDSDISQSGSVVEDDRSLQDGGASQAESKHELESTKESDEM
TSLVGNRNGDKLEHQEQENETYSNVEAASCRFPNGEIIIGATNGKNE
GEEGECFSCSRSSMEEGDATSDIGATDGNMKNENEGESCRSSMEEGDATSDISQNK

ETVEHLKXIDREKERERKQBRVMVERAIRERAPADAMERACKTAMEKAKAVA
HREVPKRSKGVENVNDLKSSAEKASQAKLAERAARVERAITETVRERAMEKALSK
SASQAKSYGGSKSPSSSGRRSSSGTENKSGSPNSNQITGEPIQRCAKASERHQ
RTSDRAAEALAEKRLDKTQKQETERNLAELADADVKKWSGKNENRALLSTLQY
ILGASGCKPIPLTDLVSSAVRKAVKATLYVHPDKLQORGASTQOKYCEKVFILL
K"
complement (9354..10899)
/gene="F22H5.16"
/note="similar to GI:1708420 from [Arabidopsis thaliana]
(J. Biol. Chem. 270 (44), 26224-26231 (1995))"
complement (join(<9354..9638,9750..9953,10043..10328,
10461..10595,10671..>10899))
/gene="F22H5.16"
complement (join(9481..9638,9750..9953,10043..10328,
10461..10595,10671..10856))
/gene="F22H5.16"
/codon_start=1
/product="NADPH oxidoreductase, putative; 10572-9197"
/protein_id="AAG12695.1"
/db_xref="GI:10092282"
/translation="MTTEKSKILVIGTGYMGEFIVEGSAKAGNPTFALVREASLSDP
VKSTIOSPKDLGVTLHGLNDHESLVKAIKQDVVISTGHKQIFDQTKIISAKE
AGNKRRLPAEFGLDVERTSAVSAFSAKGLQIRRAERAEIGPIYTVVNSAGFY
LRTLLQESGLISHTRDKAIFGDKNPPPKVTLGDKNAKVINKEEDVAAVNRKA
VDDTLNKTLYISPPNNILSMEMVTLEKIKGKLEKTHISEQILKSIQVPIDFV
KSNHAYFVKGDQTSFTIEPWFGEASVLYPDVKYTSIDYLSQFT"
complement (11122..12573)
/gene="F22H5.18"
/note="similar to GI:1708420 from [Arabidopsis thaliana]
(J. Biol. Chem. 270 (44), 26224-26231 (1995))"
complement (join(<11122..11422,11516..11734,11791..12034,
12117..12251,12333..>12573))
/gene="F22H5.18"
complement (join(11235..11422,11516..11734,11791..12034,
12117..12251,12333..12518))
/gene="F22H5.18"
/codon_start=1
/product="NADPH oxidoreductase, putative; 12234-10951"
/protein_id="AAG12680.1"
/db_xref="GI:10092267"
/translation="MASEKSKILVIGTGHIGKLIIEASVAKGHSTIALVREASLSDP
NKGKTVONFKDGVTLHGLNDHESLVKAIKQADVISTGVSMQILDQTKIISAKE
AGNKRRLPAEFGLDVERTSAVSAFSAKGLQIRRAERAEIGPIYTVVNSAGFY
LPTLVQEPGLTSPRRDKVKIFGDKNPPPKVTLGDKNAKVINKEEDVAAVNRKA
TLVYNPNTLSMNEIVTLWEKKIGKSVKEIKYSEQIFKSIQESPPVFNLLSINHA
VFVKGDQNTFTIEPWFGEASVLYPDVKYTSIDYLSYFALGTSLNT"
complement (12872..14426)
/gene="F22H5.17"
/note="similar to GI:1708420 from [Arabidopsis thaliana]
(J. Biol. Chem. 270 (44), 26224-26231 (1995))"
complement (join(<12872..13216,13298..13513,13578..13821,
13974..14108,14193..>14426))
/gene="F22H5.17"
complement (join(13053..13216,13298..13513,13578..13821,
13974..14108,14193..14378))
/gene="F22H5.17"
/codon_start=1
/product="NADPH oxidoreductase, putative; 14094-12769"
/protein_id="AAG12677.1"
/db_xref="GI:10092264"
/translation="MATEKSKILVIGTGYIGKFLVSEASAKAGHSTFALVREATLSDP
VKGTQVONFKDGVTLHGLNDHESLVKAIKQDVVISTGVSMQILDQTKIISAKE
AGNKRRLPAEFGLDVERTSAVSAFSAKGLQIRRAERAEIGPIYTVVNSAGFY
LPTLVQEPGLTSPRRDKVKIFGDKNPPPKVTLGDKNAKVINKEEDVAAVNRKA
LYIKPSNLSMNEIVTLWEKKIGKSVKEIKYSEQIFKSIQESPIPVNVLISINAV
FVNGDNTISIEPWFGEASVLYPDVKYTSIDYLSYFA"
complement (15081..16153)
/gene="F22H5.1"
/note="similar to GI:6939839 from [Oryza sativa]"
join(15081..15228,15380..15470,15661..>16153)
/gene="F22H5.1"
join(15171..15228,15380..15470,15661..16153)
/gene="F22H5.1"

gene

mRNA

CDS

gene

mRNA

CDS

gene

mRNA

CDS

gene

mRNA

CDS

```
/codon_start=1
/product="GSH-dependent dehydroascorbate reductase 1,
putative; 14887-15869"
/protein_id="AF12679.1"
/db_xref="GI:10092266"
/translation="MALDICVAVGAPVLDGCPSPQVLLTLBEKLPYKTHLINV
SKDPQFLDISPGKVPVKLDGKWADSDVIVGLEEKVPPSLKTPPEFASVGSKI
FGFVFLSKDNDGSEKALVDELEALHNLKTHSGPVAGEKITAVDLSLAPKLYH
LEVALGHYKNSVPSLTSVRNKAFLKFSRESFENTYKAKEIIVVAGWESKVN"
/gene="complement(16410..18554)"
/mRNA
/complement(join(16410..16558,16703..16830,17050..18554))
/gene="F22H5.2"
/complement(join(16410..16558,16703..16830,17050..18554))
/gene="F22H5.2"
/codon_start=1
/product="unknown protein; 18270-16126"
/protein_id="AA012682.1"
/db_xref="GI:10092269"
/translation="MPFRFRIPFFSNGPSRISSTSPSPPTPPSRPPFRPAGIA
```

Alignment Scores:

Pred. No.:	5,15e-72	Length:	68041
Score:	1117.50	Matches:	340
Percent Similarity:	38.10%	Conservative:	73
Best Local Similarity:	31.37%	Mismatches:	82
Query Match:	40.21%	Indels:	595
DB:	8	Gaps:	18

US-10-051-909-36 (1-553) x AC025814 (1-68041)

Qy	52	GlulSerGlySerAspHisAspGlyValLeuArgProLeuLeuAnthrGlySer	71
Db	33823	GAGGAGCTAGGAATGAT-----TTGGCTCGGCGGTTTATTACACCGGAAGC	33870
Qy	72	TpTYrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal	91
Db	33871	TGSPACCGGATGGGTCGACCAATCGAGTATG-----ATGGGTTCCCTCTCAAGT	33921
Qy	92	LeuArgGlnSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIle	111
Db	33922	ATTGACAGAGCTATCTCTGTCCTGTCATGCTGTTGATCGTCTCTGCGCCATTT	33981
Qy	112	GlnPheGlyPheThr-----	116
Db	33982	CAATTCGGATTCACTGTAGACAAATTCAAATCTCAACTTTTTTTTACATTTTATTC	34041
Qy	116	-----	116
Db	34042	TTGTAGACAGAGGTGACTTTTGTGTGAAACCTAGTGTGAGCTGGATTGATGATC	34101
Qy	116	-----	116
Db	34102	TTGGTTTGGTTTGATTTGTCATTTTCAATGCAAGTCTGGTGATTGAAAATCTCATAT	34161
Qy	116	-----	116
Db	34162	TGGTGAACCGAGATGATTTTCAAGTTTGTCTGACAGATTAGAGGACTTTGAAATGGGA	34221
Qy	116	-----	116
Db	34222	ATTTCTGATTTGTTGCTTAAACCTTCCATGTTTGACATTTGATCACCATCTTTAGTACT	34281
Qy	117	-----SerGlyPheSerSerProThrGlu	124
Db	34282	TGTGTTTGGTCTTTAAACCTCCATCCTTTTGTTCAGTGTGGTATTATCATCTCCAACTCA	34341
Qy	124	nAspAlaMetValArgAspLeuLeuSerIleSerGlu-----	137
Db	34342	AGCTGCAATCACCAGGATCTCGGTTTAAACAGTATCTGAGGTAATGACTGATCATTTGTC	34401
Qy	137	-----	137

Db	34402	TAGAGCTGCTGGTTTTTTTATCATTTGTTAGCGGTTTGTGATTTGATGTACCACTC	34461
Qy	138	-----PheSerAlaPheGlySerLeuSerAsnValGlyMetValGlyAlaIle	154
Db	34462	CATGTTTTAGTACTCTGTTGGTTCTCTTCCAATGTGGTGCTATGGTTGGTCAAT	34521
Qy	154	eAlaSerGlyGlnMetAlaGluTyrIleGlyValG-lysGly-----	167
Db	34522	TGCCAGTGGTTCAGATTGCAATATACATTGGACGAAAGGGGTATATCTTTGTTATCTGATT	34581
Qy	167	-----	167
Db	34582	CTTGATCGTTCTGTTTTAGATCCATTTTATAAGTCCTCTTTCTGCGTTCTAAACATCT	34641
Qy	168	-----SerLeuMetIleAlaIleProAsnIleIleGlyTyrPL	181
Db	34642	TATTGTGTTTTCTTGGACAGCTCTCTTATGATGTGCAATTTCCCAATATCATTTGATGCG	34701
Qy	181	euAlaIleSerPheAlaLys-----	187
Db	34702	TTTGATATCATTTGCG-AAAAAGTGAGTTTGTGCTCTCATACCAATAAATTTGGCGATT	34760
Qy	188	-----AspAlaSerPheL	192
Db	34761	GATTGTGTTGTTGACTTTAACTCTTTATCATCTCTGTTTGTAGGATACTCTTTTC	34820
Qy	192	euTyrMetGlyArgLeuLeuGlyPheGlyValGlyIleIleSerTyrThr-----	209
Db	34821	TTTACATGGGAAGACTTTTGAAGGCTTTGGGTTGGGATCATCTCTTACACTGTAATAC	34880
Qy	209	-----	209
Db	34881	TGAACCTTAAGTACTGTTTACTTATGGAATAACGGTTCAGTATATTAATGCAACTCGGTT	34940
Qy	210	-----ValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaL	226
Db	34941	CGTGATTTTAGTACTCTGTATATATCGTGAGATAGTCCACAAATATGAGAGGAGAT	35000
Qy	226	euGlySerValAsnGlnLeu-----	232
Db	35001	TGGGTTTCGGTTAAACAGGT-TAGACATGTAGTAGTACTAATTCCTGTTTGTGCTTTGTG	35059
Qy	233	-----SerValT	235
Db	35060	TACTGACTGTTAATGCTAAACAGAAATGTTGTGTAACCTTTTGTGCTGTAGTCTTCTGTA	35119
Qy	235	hrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProThrArgLeuLeuAlaV	255
Db	35120	CAATTGGAATAATGTGGCGTATTTACTTTGCTTGTGCTTGTCCATGGAGAAATCTTGCGAG	35179
Qy	255	al-----	255
Db	35180	TTCTGGTAAATCACTATTACACGGCTGAGACTATTTTGTGATTATGAATAATTTGCTAC	35239
Qy	256	-----IleGlyAlaLeuProCysThrMetL	264
Db	35240	GCTCTAACTCCTTCTCTGTTCTCTTCTCTTTATCTTTAGGATATTGGCGTGTACATAT	35299
Qy	264	euIleProGlyLeuPhePheIleProGluSerProArgTyrIleu-----	278
Db	35300	TGATACACAGGTCTCTTTTTCATCCCGAATCCCTCGCTCGTGGTGGTAAAGCCTTACAGTAG	35359
Qy	278	-----	278
Db	35360	TTTGCTTATCTTCTGAGTTTAGTGCAATTTTGGTTACTCCACTCTTTTTTTTTTTTT	35419
Qy	279	-----AlaLysM	281
Db	35420	TAATAACATCAATTAATCATCTCTGTGGGATACTTGAATTCGTTAAAAATGTAGGCAAAA	35479
Qy	281	etAsnLeuThrGluAspCysGluThrSerLeuGlnValIleuArgGlyPheGluThrAspI	301
Db	35480	TGGGTATGACAGATGAATTTGAAATCTTATTACAAAGTTCTTCCAGATTGTGACTGATA	35539

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 13, 2004, 11:39:19 ; Search time 676.434 Seconds

(Without alignments)
4291.521 Million cell updates/sec

Title: US-10-051-909-36

Perfect score: 2779
Sequence: 1 PSSSSFRPAGKKKKKNGQ.....TIFVSLSIQRQLQWLPECLUS 553

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool.h/US10051909/runat_13102004_123335_19900/app_query.fasta_1.1678
-DB=N Geneseq 23Sep04 -QFMT=fastap -SUFFIX=mg -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051909.ecgn_1_891_@runat_13102004_123335_19900 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCKS=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: Geneseq_23Sep04.*
2: Geneseqn1980s.*
3: Geneseqn1990s.*
4: Geneseqn2000s.*
5: Geneseqn2001as.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2777	99.9	1953	12	ADG47939 Corn Beta
2	1710	61.5	1377	8	ADA69666 Rice gene
3	1577	56.7	1428	8	ADA69667 Rice gene
4	951.5	34.2	1741	3	AAC36319 Arabidops
5	943.5	34.0	1683	3	AAC39099 Arabidops
6	933	33.6	1395	3	AAC45857 Arabidops

ALIGNMENTS

RESULT 1	ADG47939	standard; cDNA; 1953 BP.	
ID	ADG47939	standard; cDNA; 1953 BP.	
XX	AC	ADG47939;	
XX	DT	11-MAR-2004 (first entry)	
XX	DE	Corn Beta-vulgaris-like sugar transport protein cDNA #4.	
XX	KW	Arabidopsis-like sugar transport protein;	
XX	KW	Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;	
XX	KW	corn; plant; gene; ss.	
XX	OS	Zea mays.	
XX	Key	Location/Qualifiers	
FT	FT	1..1662	
FT	FT	/tag= a	
FT	FT	/product= "Corn Beta-vulgaris-like sugar transport	
FT	FT	protein"	
FT	FT	/note= "Xaa can be any amino acid, CDS does not include	
FT	FT	no start codon"	
FT	FT	/transl_except= {pos:1582..1584, aa:Xaa}	
FT	FT	/partial	
XX	XX	US2002199217-A1.	
XX	PD	26-DEC-2002.	

7	923	33.2	1685	3	AAC45298	Arabidops
8	923	33.2	1743	3	AAC45290	Arabidops
9	920.5	33.1	1473	6	ABZ14526	Arabidops
10	916	33.0	1755	3	AAC40459	Arabidops
11	899	32.3	1359	6	ABZ13822	Arabidops
12	893	32.1	1768	12	ADMA7939	Polynucle
13	877.5	31.6	1407	6	ABZ13486	Arabidops
14	877.5	31.6	1407	8	ADA67841	Arabidops
15	877.5	31.6	1642	3	AAC46009	Arabidops
16	868.5	31.3	1425	3	AAC50050	Arabidops
17	853.5	30.7	1425	3	AAC50049	Arabidops
18	837.5	30.1	1209	8	ADA70113	Rice gene
19	824	29.7	1452	3	AAC46559	Arabidops
20	805.5	29.0	1695	3	AAC49517	Arabidops
21	790.5	28.4	1436	3	AAC45400	Arabidops
22	786.5	28.3	1440	3	AAC33666	Arabidops
23	777	28.0	1374	6	ABZ13215	Arabidops
24	758.5	27.3	1378	3	AAC35123	Arabidops
25	754.5	27.2	1344	8	ADA69777	Rice gene
26	688.5	24.8	1724	4	ABZ13389	Drosophil
27	682.5	24.6	1239	4	AAC34742	Arabidops
28	674.5	24.3	2072	4	AAF55867	Murine GLU
29	674	24.3	2217	4	AAF55865	Human GLU
30	672	24.2	898	10	ADD17645	Ad17645 DNA (Seq1
31	672	24.2	2087	4	AAF55866	Rat GLUTX
32	669.5	24.1	2080	4	AD09552	Human tra
33	663	23.9	1473	6	ABZ13864	Arabidops
34	663	23.9	1473	6	ADG88039	A. thalia
35	658.5	23.7	1445	8	ABZ24794	Human sol
36	658.5	23.7	1856	8	ABZ24792	Human sol
37	652.5	23.5	1873	8	ABZ24793	Human sol
38	652.5	23.5	2514	4	ABL17067	Drosophil
39	640	23.0	1412	4	ABZ13883	Drosophil
40	623.5	22.4	2487	4	ABZ24790	Human sol
41	621.5	22.4	1541	4	AAF55870	Human GLU
42	621	22.3	2356	8	ADA84076	Human SUC
43	621	22.3	2356	8	ABZ24791	Human sol
44	621	22.3	2390	4	AAK94338	Human ful
45	621	22.3	2390	12	ADL31000	Full leng

Db 1621 ATTACAGCTCAGCTGCATGCGTTGCCGAGTGTTTACT 1659

RESULT 2
ADA69666
ADA69666 standard; DNA; 1377 BP.
XX AC
XX ADA69666;
XX DT 20-NOV-2003 (first entry)
XX XX
XX Rice gene, SEQ ID 2989.
XX DE
XX Plant; bacterial infection; fungal infection; viral infection; rice,
KW gene; ds.
XX KW
XX Oryza sativa.
XX OS
XX WO2003000898-A1.
XX FN
XX 03-JAN-2003.
XX PD
XX XX
XX 22-JUN-2001; 2001WO-IB001105.
XX PF
XX 22-JUN-2001; 2001WO-IB001105.
XX PR
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX PA
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Qian S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
PI
XX WPI; 2003-175290/17.
XX DR
XX Identifying at least one gene involved in plant resistance or response to
XX pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX PS
XX Claim 6; SEQ ID NO 2989; 899pp; English.
XX CC
XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC the expression of the gene in an uninfected plant in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.

SQ Sequence 1377 BP; 290 A; 338 C; 360 G; 388 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: Length: 1377
Score: 1710.00 Matches: 358
Percent Similarity: 78.36% Conservative: 44
Best Local Similarity: 69.79% Mismatches: 39
Query Match: 61.53% Indels: 73
DB: Gaps: 9

US-10-051-909-36 (1-553) x ADA69666 (1-1377)

QY 44 AsnArg--GlyGlyAlaGlyAlaGlyGluSerGlySerAspHisAsp-----Gly 60
DB 4 AACAGGGCGGCGCGCTGCGGCTCGCGATGAGCGCGAGCGACTACGAGAGCGCGCG 63
QY 61 ValLeuArgArgProLeuLeu---AenThrGlySerTrpTyArgMetSerSerArgGln 79
DB 64 GC CATGCGGAAGCGCGCTGTGTGATGCACACGGGAGCGTGTACAGCATGGGTTCGCGCAG 123
QY 80 SerSerPhe---AlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValser 98

QY 447 ileThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuPro 466
 DB 1099 ATCGCTTCTCTTCGGTATGGTGCCATTCATGGATCAATAGTCAGAGATCTCCG 1158
 QY 467 ValSerIleLysSerLeuGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPhe 486
 DB 1159 GTTAGATCAAGAGTCTCGAGGAAGCTTTCGACGCTCGCCAACTGGCTTACATCTTT 1218
 QY 487 AlaIleThrMetThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSer 506
 DB 1219 GGAATAACAATGACAGCAACTTCATGCTTAGCTGGAGTCTGGAGGACCTTTGTGTC 1278
 QY 507 TyrMetValValSerAlaPheThrIleValPheValValLeuTrpValProGluThrLys 526
 DB 1279 TACATGGTGTGAGTGTCTTTCACCTCGTGTCTCATCTCTTGGGTGCCAGACAAAA 1338
 QY 527 Gly***AsnSerArgGlyAspThrIlePheValSerLeu 539
 DB 1339 GGAAG-AACTCTGAAGAGATACATGTCTCTCCGCTG 1376

RESULT 3

ADA69667
 ID ADA69667 standard; DNA; 1428 BP.

AC ADA69667;

DT 20-NOV-2003 (first entry)

XX Rice gene, SEQ ID 2990.

DE plant; bacterial infection; fungal infection; viral infection; rice;
 KW gene; ds.

XX Oryza sativa.

XX WO2003000898-A1.

XX 03-JAN-2003.

XX 22-JUN-2001; 2001WO-IB001105.

XX 22-JUN-2001; 2001WO-IB001105.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting a plant
 PT gene expression.

XX Claim 6; SEQ ID NO 2990; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance to a response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.

XX SQ Sequence 1428 BP; 352 A; 315 C; 354 G; 407 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.42e-133 Length: 1428

Score: 1577.00 Matches: 318
 Percent Similarity: 85.85% Conservative: 34
 Best Local Similarity: 77.56% Mismatches: 32
 Query Match: 56.75% Indels: 27
 DB: 8 Gaps: 2

US-10-051-909-36 (1-553) x ADA69667 (1-1428)

QY 130 AspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGly 149
 DB 277 GATCTGTACTTACAACAGCATATGTTCTCGGTTTTCGGTTCCTCTCCAAAGTCGGCCGC 336
 QY 150 MetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyValGlySerLeu 169
 DB 337 ATGGTCGGAGCGATCGCCAGTGGCGAGATGGCGAGTACATTGGCCGGAAGGTCGTTG 396
 QY 170 MetIleAlaIleAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAla 189
 DB 397 ATAATTGGCGGCTTCTTAACATCAATTGGTTGGCTTGCCATCTCTTTGCAAAAGACGCG 456
 QY 190 SerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThr 209
 DB 457 TCATTTTATACATGGGACGCTTCTTGAAGTTTTCGGTGTGGTTATATCATATACG 516
 QY 210 ValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerVal 229
 DB 517 GTGCCAGTATACATAGCAGAGATATCTCATCAGAACACACAGAGGAGCACTTGGCTCGGTG 576
 QY 230 AsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIlePro 249
 DB 577 AACCAAGTTGTCGTTACCATTTGGTATCTTGTGGCTATTTGGTATTTGGTATTTGGTCT 636
 QY 250 TrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhe 269
 DB 637 TGGAGGCTGCTTGCAGTGATAGGAAGCATCCATGATCATTTGTTTAACTGGTCTATTTC 696
 QY 270 PheIleProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThr 289
 DB 697 TTCATTCGGAATCCCAAGATGGCTGGCAAAATGAAATGATGATGATGATGATGATGATG 756
 QY 290 SerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLys 309
 DB 757 TCTCTACAAGTTTGGAGGGGTTTGAGACTGATATTACTGCAGAA----- 801
 QY 310 ArgAlaValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLys 329
 DB 802 AGACAGTAGCATCGGCAAAACAAAGAACACACAGTCCGCTTTTAAAGAGATTGAACCAAAAG 861
 QY 330 LysTyrArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGly 349
 DB 862 AAATACCGCATCTCTTACTGATAGGAAGTGGCTTCTTCTACTTCAGATCTAAGTGA 921
 QY 350 IleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSer 369
 DB 922 ATAAATGGTATCTGTTTATGCAAGTAGAATCTTCAGAGATGACGGGTTTACAAACAGT 981
 QY 370 AspLeuAlaThrCysSerLeuGlyAlaIleGlyValLeuAlaThrGlyValThrTrp 389
 DB 982 GACTTGGCCACATGTGCATAGGAGCAATTGAG----- 1014
 QY 390 LeuLeuAspArgAlaGlyArgA-gIleLeuLeuIleIleSerThrSerGlyMetThrLeu 409
 DB 1015 -----ATCTACTGCTGGGATGACTCTA 1038
 QY 410 CysLeuLeuAlaValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsn 429
 DB 1039 AGCTTCTTCAGTATCTGTGTATTCTTCTGGAGGGTAATATTTCACATGATTCTCAT 1098
 QY 430 SerTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPhe 449
 DB 1099 TCGTTCTACATCTTAAGTATGATCTCTTGGTGTGCTTGTGGCTTATATCATCATCTTT 1158
 QY 450 SerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIle 469

Db	1159	TCCTTCGGCATGGTCCCATTCATGGGTTCATGATCTCTGAGATCCTCCCATTTAGCATC	1218	99US-0136021P
				99US-0136382P
				99US-0136782P
Qy	470	LyeSerLeuClVgLySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThr	489	99US-0137222P
				99US-0137528P
Db	1219	AAGAGTCITTTGGGGGAAGCTTTGGCAGACCTTGGCAACATCTTACATCTCTGGGCAATACCA	1278	99US-0137502P
				99US-0137724P
Qy	490	MetThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetVal	509	99US-0138094P
				99US-0138540P
Db	1279	ATGACAGCAAAATTTGTTACTTCAGCTGGAGTGTCTGGAGGAACAATTTTTTGTCTACATGATT	1338	99US-0138847P
				99US-0139119P
Qy	510	ValSerAlaPheThrIleValPheValLeuTrpValProGluThrIysGly**Asn	529	99US-0139453P
				99US-0139453P
Db	1339	GTGAGTGCITTCACCTCTCGTGITTCGTCATATTTTGGTGCAGAGCAAGGGAAG-AAC	1397	99US-0139454P
				99US-0139454P
Qy	530	SerArgGlyAspThrIlePheValSerLeu	539	99US-0139455P
				99US-0139456P
Db	1398	TCTGGAGGACATACAGTTTCGTTCTGCTA	1427	99US-0139457P
				99US-0139458P
RESULT 4				
AAC36319				
ID AAC36319 standard; DNA; 1741 BP.				
XX	AC	AAC36319;		
XX	AC	AAC36319;		
XX	AC	AAC36319;		
DT	17-OCT-2000	(first entry)		
XX	XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 13356.		
DE	XX	Arabidopsis thaliana		
XX	XX	Hybridisation assay; genetic mapping; gene expression control;		
KW	KW	protein identification; signal transduction pathway; metabolic pathway;		
KW	KW	promoter; termination sequence; ss.		
XX	XX	Arabidopsis thaliana.		
OS	XX	Arabidopsis thaliana.		
XX	XX	EP1033405-A2.		
PN	XX	06-SEP-2000.		
XX	XX	06-SEP-2000.		
XX	XX	25-FEB-2000; 2000EP-00301439.		
XX	XX	25-FEB-1999; 99US-0121825P.		
PR	PR	05-MAR-1999; 99US-0123180P.		
PR	PR	09-MAR-1999; 99US-0123548P.		
PR	PR	23-MAR-1999; 99US-0125788P.		
PR	PR	25-MAR-1999; 99US-0126264P.		
PR	PR	29-MAR-1999; 99US-0126785P.		
PR	PR	01-APR-1999; 99US-0127482P.		
PR	PR	06-APR-1999; 99US-01282234P.		
PR	PR	08-APR-1999; 99US-0128714P.		
PR	PR	16-APR-1999; 99US-0129845P.		
PR	PR	19-APR-1999; 99US-0130077P.		
PR	PR	21-APR-1999; 99US-0130449P.		
PR	PR	23-APR-1999; 99US-0130510P.		
PR	PR	23-APR-1999; 99US-0130881P.		
PR	PR	28-APR-1999; 99US-0131449P.		
PR	PR	30-APR-1999; 99US-0132048P.		
PR	PR	04-MAY-1999; 99US-0132407P.		
PR	PR	05-MAY-1999; 99US-0132484P.		
PR	PR	05-MAY-1999; 99US-0132485P.		
PR	PR	06-MAY-1999; 99US-0132486P.		
PR	PR	06-MAY-1999; 99US-0132487P.		
PR	PR	07-MAY-1999; 99US-0132863P.		
PR	PR	11-MAY-1999; 99US-0134256P.		
PR	PR	14-MAY-1999; 99US-0134218P.		
PR	PR	14-MAY-1999; 99US-0134219P.		
PR	PR	14-MAY-1999; 99US-0134221P.</		

```

PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149446P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149939P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 04-OCT-1999; 99US-0157553P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159233P.
PR 13-OCT-1999; 99US-0159234P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161520P.
PR 28-OCT-1999; 99US-0161592P.

PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 2,528-76 Length: 1741
Score: 951.50 Matches: 181
Percent Similarity: 66.12% Conservative: 102
Best Local Similarity: 42.23% Mismatches: 140
Query Match: 34:24% Indels: 5
DB: 3 Gaps: 2

US-10-051-909-36 (1-553) x AAC36319 (1-1741)
QY 100 PheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPhe 119
Db 287 TATCTCAGCACATTCGTTGCGGTCTGTGGTCTCTTTTGGGTTCTGTTGCGGGATAC 346
QY 120 SerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSer 139
Db 347 TCTTCACCTGCTCAGCTGCAATAGGAATGACCTTTTCATTGACTATAGCTGAGTTTCA 406
QY 140 AlaPheGlySerLeuSerAsnValGlyMetValGlyAlaIleAlaSerGlyGlnMet 159
Db 407 CTCCTTTGGTTCTTACTAATCTTTGGCGCAATGATCGGTCTATACAAAGCGGCCATATA 466
QY 160 AlaGluTyrIleGlyArgGlySerLeuMetIleAlaAlaIleProAsnIleIleGly 179
Db 467 GCTGATTTAGTTGGAAGAAAGGGGCGATGAGATTCTCTCGCATTTTGTGTAGTCGGG 526
QY 180 TrpLeuAlaIleSerPheAlaIleAspAlaSerPheLeuTyrMetGlyArgLeuGlu 199
Db 527 TGGTAGCAATCACTTTGCCAAGGGAGTAGTGCTCTGGACCTCGGAAGACTGGCAACG 586
QY 200 GlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPro 219
Db 587 GGATATGAAATGGGAGCATTTTCTATGTGTGCAATCTTTATAGCAGAAATATGACCT 646
QY 220 GlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePhe 239
Db 647 AAACCTTCAGAGGGCTCTAACCACTGACAGATTCTGATCTGCATCTGAGTGTGCG 706
QY 240 LeuAlaTyrLeuLeuGlyMetPheIleProTrpArgLeuLeuAlaIleGlyAlaLeu 259
Db 707 GTTTCCTTCATCATAGGCACACTAGTCAGCTGGAGAGCTTGGCATTATAGGAATCATC 766
QY 260 ProCysThrMetLeuIleProGlyLeuPheIleProGlySerProArgTrpLeuAla 279
Db 767 CCATGCGCTGCTCTCTCTTGGCTCTTTTATCCCTGAGTCTCCAAAGATGCTGGCA 826
QY 280 LysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThr 299
Db 827 AAAGTGGGGCTGATACCGAGTTTGAAGCTGCATCGAGAAAGCTCCGTGGGAAGAGCT 886
QY 300 AspIleThrThrGluValAsnAspIleIleArgAlaValAlaSerSerSerLysArgThr 319
Db 887 GATATTTGGAGGAGGAGCAGAGATCCAGGATTATATCGAAACTCTGGAAGAGCTACCG 946
QY 320 ThrIleSerPheGlnGluLeuAsnGlnLysTyrArgThrProLeuLeuLeuGlyIle 339
Db 947 AAAGCCAGATGCTGGATTGTTTCAGAGGAGATACATACGCTCTGTCTTATAGCTTTC 1006
QY 340 GlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSer 359
Db 1007 GGGTTGATGTTGTTTTCAGCAGATTTGGAGGAATCAACGGAATATGTTTTCACACAGCTCG 1066
QY 360 IlePheLysAlaIleGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIle 379
Db 1067 ATATTGAGCAAGCAGGTTTTCCTCC--ACAAGACTTGGGATGATATATATGCTGTCTT 1123
QY 380 GlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspAlaGlyArgArgIleLeu 399
Db 1124 CAGGTGGTAATCACTGCGCTTAATGACCCGATAGTTGACAGAGCCGGAAGAAACCATTTG 1183

```

Qy	400	LeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValPhePhe	419
Db	1184	CTACTGTGTTCTCGCAACAGGGTTAGTGATAGCCTGTTTGGATAGCAGCGGTTTCCTTCTCAT	1243
Qy	420	LeuLysAspAsnIleSerGlnApsSerAsnSerTyrtyrileLeuThrMetIleSerLeu	439
Db	1244	CTCAGGTTTCAC-----GACATGGCGCACGAAGCATCCAGTCCCTGGCTGTT	1291
Qy	440	ValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeu	459
Db	1292	GTTGGTAATAATGGTGATACATAGGATCGTTTTACGACGAATGGAGCAATGCCGTGGGTG	1351
Qy	460	MetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeu	479
Db	1352	GTCATGCTGAGATATTCCCATAAACATAAAGGATGACGAGGCATGGCGACCGCTG	1411
Qy	480	AlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSer	499
Db	1412	GTGAACGTGTTTGGAGCGTGGCTGTTCTTACACTTTCAAATTCCTCATGTCTCTGGAGC	1471
Qy	500	ValGlyGlyThrPheLeuSerTyrtMetValValSerAlaPheThrIleValPheValVal	519
Db	1472	TCTTACGGAACTTTCCTCATTTACGTGCGCATCAACGCGTGGCCATTGTTTGGTCATT	1531
Qy	520	LeuTrpValProGluThrLysGly	527
Db	1532	GCCATCGTGCCTGAGACAAAAGG	1555
<hr/>			
RESULT 5			
AAC39099			
ID	AAC39099	standard; DNA; 1683 BP.	
XX	AC	AAC39099;	
XX	XX		
DT	17-OCT-2000	(first entry)	
DE	XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 23364.	
KW	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic pathway;		
KW	promoter; termination sequence; ss.		
OS	Arabidopsis thaliana.		
XX	EP1033405-A2.		
PN	06-SEP-2000.		
PD	XX		
PF	25-FEB-2000; 2000EP-00301439.		
XX	XX		
PR	25-FEB-1999;	99US-0121825P.	
PR	09-MAR-1999;	99US-0123180P.	
PR	09-MAR-1999;	99US-0123548P.	
PR	23-MAR-1999;	99US-0125788P.	
PR	25-MAR-1999;	99US-0126264P.	
PR	23-MAR-1999;	99US-0126785P.	
PR	01-APR-1999;	99US-0127462P.	
PR	06-APR-1999;	99US-0128234P.	
PR	08-APR-1999;	99US-0128714P.	
PR	16-APR-1999;	99US-0129845P.	
PR	19-APR-1999;	99US-0130077P.	
PR	21-APR-1999;	99US-0130449P.	
PR	23-APR-1999;	99US-0130510P.	
PR	23-APR-1999;	99US-0130891P.	
PR	28-APR-1999;	99US-0131449P.	
PR	30-APR-1999;	99US-0132048P.	
PR	30-APR-1999;	99US-0132407P.	
PR	04-MAY-1999;	99US-0132484P.	
PR	05-MAY-1999;	99US-0132485P.	
PR	06-MAY-1999;	99US-0132486P.	
PR	07-MAY-1999;	99US-0132487P.	
PR	11-MAY-1999;	99US-0132863P.	
PR	11-MAY-1999;	99US-0134256P.	

PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161922P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161933P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.	Alignment Scores:		
PR	05-AUG-1999;	99US-0147192P.	Pred. No.:		
PR	05-AUG-1999;	99US-0147260P.	Score:		
PR	06-AUG-1999;	99US-0147303P.	Percent Similarity:		
PR	06-AUG-1999;	99US-0147416P.	Best Local Similarity:		
PR	09-AUG-1999;	99US-0147493P.	Query Match:		
PR	09-AUG-1999;	99US-0147935P.	DB:		
PR	10-AUG-1999;	99US-0148171P.	Length:		
PR	11-AUG-1999;	99US-0148319P.	Matches:		
PR	12-AUG-1999;	99US-0148341P.	Conservative:		
PR	13-AUG-1999;	99US-0148565P.	Mismatches:		
PR	13-AUG-1999;	99US-0148684P.	Indels:		
PR	16-AUG-1999;	99US-0149175P.	Gaps:		
PR	17-AUG-1999;	99US-0149175P.	US-10-051-909-36 (1-553) x AAC39099 (1-1683)		
PR	18-AUG-1999;	99US-0149426P.	QY	100 PheLeuCysThrLeuLeuValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPhe	119
PR	20-AUG-1999;	99US-0149722P.	DB	207 TATCTCAGCACATTCGTTGCCGTCGTGCTTCTTGGCTTCGGTCTTGTGGCGGATAC	266
PR	20-AUG-1999;	99US-0149723P.	QY	120 SerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGlyPheSer	139
PR	20-AUG-1999;	99US-0149923P.	DB	267 TCTTCACCTGCTCAGGCTGCAATAGGAATGACCTTTTCATTGACTATAGCTGAGTTTCA	326
PR	23-AUG-1999;	99US-0149902P.	QY	140 AlaPheGlySerLeuSerAsnValGlyMetValGlyAlaIleAlaSerGlyGlnMet	159
PR	25-AUG-1999;	99US-0150566P.	DB	327 CTCCTTGGTCTTCTACTAACTTTTGGCGCAATGATCGGTCTGTAAACAGCGGCCTATA	386
PR	26-AUG-1999;	99US-0150884P.	QY	160 AlaGluTyrIleGlyValGlyValGlySerLeuMetIleAlaAlaIleProAsnIleIleGly	179
PR	27-AUG-1999;	99US-0151065P.	DB	387 GCTGATTTAGTTGGAGAAAGGGGGGATGAGAGTTCTCTGCATTTTGTGTAGTCGG	446
PR	27-AUG-1999;	99US-0151086P.	QY	180 TrpLeuAlaIleSerPheAlaIleAspAlaSerPheLeuTyrMetGlyArgLeuLeuGlu	199
PR	31-AUG-1999;	99US-0151303P.	DB	447 TGGCTAGCAATCATCTTTGCCAAGGGAGTAGTGGCTCTGGACCTCGGAAGACTGGCAACG	506
PR	07-SEP-1999;	99US-0151930P.	QY	200 GlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPro	219
PR	10-SEP-1999;	99US-0153070P.	DB	507 GGATATGAATGGGAGCATTTCTCTATGCTGCGCAATCTTTATAGCAATATTCACCT	566
PR	13-SEP-1999;	99US-0153758P.	QY	220 GlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePhe	239
PR	16-SEP-1999;	99US-0154039P.	DB	567 AAACTTTTCAGAGGGGCTCTAACCACTGAACCAAGATTCTGATCTGCACCTGGAGTGTG	626
PR	22-SEP-1999;	99US-0155139P.	QY	240 LeuAlaTyrLeuLeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeu	259
PR	23-SEP-1999;	99US-0155486P.	DB	627 GTTCTCTTCATCATTAGGCACACTAGTAGCTGGAGAGTCTTGGCATTAATAGGAATCATC	686
PR	24-SEP-1999;	99US-0156599P.	QY	260 ProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeuAla	279
PR	28-SEP-1999;	99US-0156458P.	DB	687 CCAATGGCTGCTCCTCTCTTGGCCTCTTTTATCTCCCTGAGTCTCCAAAGTGGTGCA	746
PR	29-SEP-1999;	99US-0156596P.	QY	280 LysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuAtgGlyPheGluThr	299
PR	04-OCT-1999;	99US-0157117P.	DB	747 AAAGTGGGGCGGTGATACCGGAGTTGAAGCTGCATCGAGGAGCTCCCGGGGAAGAGCT	806
PR	05-OCT-1999;	99US-0157753P.	QY	300 AspIleThrThrGluValAsnAspIleLysArgAlaValAlaLaserSerSerLysArgThr	319
PR	06-OCT-1999;	99US-0157865P.	DB	807 GATATTCGGAGGAGCAGCAGAGATCCAGGATTATATCGAACTCTCGAAAGGCTACCG	866
PR	07-OCT-1999;	99US-0158029P.	QY	320 ThrIleSerPheGlnGluLeuAsnGlnLysTyrArgThrProLeuLeuLeuGlyIle	339
PR	08-OCT-1999;	99US-0158232P.	DB	867 AAAGCCAAAGATGCTGGATTGTTTTCAGAGGAGATACATACCTCTGTCTTATAGCTTTC	926
PR	12-OCT-1999;	99US-0158389P.	QY	340 GlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSer	359
PR	13-OCT-1999;	99US-0159253P.	DB	927 GGGTTGATGGTGTTCAGCAGTTGGAGGATCAACGAATATGTTTCTACACAGCTCG	986
PR	13-OCT-1999;	99US-0159294P.			
PR	13-OCT-1999;	99US-0159295P.			
PR	14-OCT-1999;	99US-0159329P.			
PR	14-OCT-1999;	99US-0159330P.			
PR	14-OCT-1999;	99US-0159331P.			
PR	14-OCT-1999;	99US-0159637P.			
PR	18-OCT-1999;	99US-0159584P.			
PR	21-OCT-1999;	99US-0160741P.			
PR	21-OCT-1999;	99US-0160767P.			
PR	21-OCT-1999;	99US-0160768P.			
PR	21-OCT-1999;	99US-0160770P.			
PR	21-OCT-1999;	99US-0160814P.			
PR	21-OCT-1999;	99US-0160845P.			
PR	22-OCT-1999;	99US-0160980P.			
PR	22-OCT-1999;	99US-0160981P.			

PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145193P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145219P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145291P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145291P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145291P.	PR	25-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145291P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146338P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146338P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146338P.	PR	28-OCT-1999;	99US-0161922P.
PR	02-AUG-1999;	99US-0146338P.	PR	28-OCT-1999;	99US-0161933P.
PR	03-AUG-1999;	99US-0147038P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151930P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
PR	22-SEP-1999;	99US-0155139P.			
PR	23-SEP-1999;	99US-0155486P.			
PR	24-SEP-1999;	99US-0155659P.			
PR	28-SEP-1999;	99US-0156458P.			
PR	29-SEP-1999;	99US-0156596P.			
PR	04-OCT-1999;	99US-0157117P.			
PR	05-OCT-1999;	99US-0157753P.			
PR	06-OCT-1999;	99US-0157865P.			
PR	07-OCT-1999;	99US-0158029P.			
PR	08-OCT-1999;	99US-0158232P.			
PR	12-OCT-1999;	99US-0158349P.			
PR	13-OCT-1999;	99US-0159283P.			
PR	13-OCT-1999;	99US-0159294P.			
PR	13-OCT-1999;	99US-0159295P.			
PR	14-OCT-1999;	99US-0159329P.			
PR	14-OCT-1999;	99US-0159330P.			
PR	14-OCT-1999;	99US-0159331P.			
PR	14-OCT-1999;	99US-0159637P.			
PR	14-OCT-1999;	99US-0159638P.			

Alignment Scores:
 Pred. No.: 8,77e-75
 Score: 933.00
 Percent Similarity: 65.73%
 Best Local Similarity: 41.96%
 Query Match: 33.57%
 DB: 3
 Length: 1395
 Matches: 180
 Conservative: 102
 Mismatches: 141
 Indels: 6
 Gaps: 3

US-10-051-909-36 (1-553) x AAC45857 (1-1395)

QY	100	PheLeuCysThrLeuLeuValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPhe	119
DB	85	TATCTCAGCACATTCGTTGGCGCTGCTGCTTTTGGCTTCGGTTCTTGTGCGGGATAC	144
QY	120	SerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSer	139
DB	145	TCTTCACCTGCTCAGGCTGCAATAGGAATCACCTTTTCATTGATAGCTAGCTGATTTCA	204
QY	140	AlaPheGlySerLeuSerAsnValGlyMetValGlyAlaIleAlaSerGlyGlnMet	159
DB	205	CTCTTTGGTTCTTACTAATCTTTGGCGCAATGATCGGTGCTATACACAGCGGCCTATA	264
QY	160	AlaGluTyrlleGlyArgGlySerLeuMetIleAlaAlaIleProAsnIleIleGly	179
DB	265	GCTGATTAGTTGGAGAGGGGGCGATGAGAGTTCTCTGCAATTTTGTGTAGTCGG	324
QY	180	TyrLeuAlaIleSerPheAlaIlyAspAlaSerPheLeuTyMetGlyArgLeuLeuGlu	199
DB	325	TGGCTAGCAATCATCTTTGCCAAGGGAGTAGTGCTCTGGACCTTGGAGACTGGCAACG	384
QY	200	GlyPheGlyValGlyIleIleSerTyThrValProValTyrlleAlaGluIleSerPro	219
DB	385	GGATATGGAATGGAGCATTTCTTATGTGGTGCATCTTTATAGCAGAAATGCACCT	444
QY	220	GlnAsnMetArgGlyAlaLeuGlySerValAsn---GlnLeuSerValThrPheGlyIle	238
DB	445	AAAACCTTTTCAGAGGGGCTCTAACACACTGAACACAGCAGATTCTGATCTGCACCTGGAGTG	504
QY	239	PheLeuAlaTyrlleLeuGlyMetPheIleProIleProIleProIleProIleGlyAla	258
DB	505	TCGGTTCTTCATCATAGGCACACTAGTAGCGTAGGAGAGCTTGGCATTATAGGATC	564
QY	259	LeuProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTyrIleu	278
DB	565	ATCCCATGGCTGCTCTCTCTTGGCTCTTTTATCCCTGAGTCTCCAGATGGCTG	624
QY	279	AlaIlyMetAsnLeuThrGluAspCysGluThrSerIleGlnValLeuArgGlyPheGlu	298
DB	625	GCAAAAGTGGGCGGTGATACGAGTTTGAAGCTGACCTGAGAGAGCTCCGTTGGGAAGAG	684
QY	299	ThrAspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerSerIlyArg	318

19-JUL-1999; 99US-0144325P. PR
19-JUL-1999; 99US-0144331P. PR
19-JUL-1999; 99US-0144332P. PR
19-JUL-1999; 99US-0144333P. PR
19-JUL-1999; 99US-0144334P. PR
19-JUL-1999; 99US-0144335P. PR
20-JUL-1999; 99US-0144352P. PR
20-JUL-1999; 99US-0144632P. PR
20-JUL-1999; 99US-0144884P. PR
21-JUL-1999; 99US-0144814P. PR
21-JUL-1999; 99US-0145086P. PR
21-JUL-1999; 99US-0145088P. PR
22-JUL-1999; 99US-0145085P. PR
22-JUL-1999; 99US-0145087P. PR
22-JUL-1999; 99US-0145089P. PR
23-JUL-1999; 99US-0145145P. PR
23-JUL-1999; 99US-0145218P. PR
23-JUL-1999; 99US-0145224P. PR
26-JUL-1999; 99US-0145278P. PR
27-JUL-1999; 99US-0145913P. PR
27-JUL-1999; 99US-0145918P. PR
27-JUL-1999; 99US-0145919P. PR
28-JUL-1999; 99US-0145951P. PR
02-AUG-1999; 99US-0146386P. PR
02-AUG-1999; 99US-0146388P. PR
02-AUG-1999; 99US-0146389P. PR
03-AUG-1999; 99US-0147038P. PR
04-AUG-1999; 99US-0147204P. PR
04-AUG-1999; 99US-0147302P. PR
05-AUG-1999; 99US-0147192P. PR
05-AUG-1999; 99US-0147260P. PR
06-AUG-1999; 99US-0147303P. PR
06-AUG-1999; 99US-0147416P. PR
09-AUG-1999; 99US-0147493P. PR
09-AUG-1999; 99US-0149175P. PR
10-AUG-1999; 99US-0148317P. PR
11-AUG-1999; 99US-0148319P. PR
12-AUG-1999; 99US-0148341P. PR
13-AUG-1999; 99US-0148565P. PR
13-AUG-1999; 99US-0148684P. PR
16-AUG-1999; 99US-0149368P. PR
17-AUG-1999; 99US-0149175P. PR
18-AUG-1999; 99US-0149426P. PR
20-AUG-1999; 99US-0149722P. PR
20-AUG-1999; 99US-0149723P. PR
20-AUG-1999; 99US-0149823P. PR
23-AUG-1999; 99US-0149902P. PR
23-AUG-1999; 99US-0149930P. PR
25-AUG-1999; 99US-0150566P. PR
26-AUG-1999; 99US-0150884P. PR
27-AUG-1999; 99US-0151065P. PR
27-AUG-1999; 99US-0151066P. PR
27-AUG-1999; 99US-0151080P. PR
30-AUG-1999; 99US-0151308P. PR
31-AUG-1999; 99US-0151338P. PR
01-SEP-1999; 99US-0151930P. PR
07-SEP-1999; 99US-0152363P. PR
10-SEP-1999; 99US-0153070P. PR
13-SEP-1999; 99US-0153758P. PR
15-SEP-1999; 99US-0154018P. PR
16-SEP-1999; 99US-0154039P. PR
20-SEP-1999; 99US-0154779P. PR
22-SEP-1999; 99US-0155139P. PR
23-SEP-1999; 99US-0155486P. PR
24-SEP-1999; 99US-0155659P. PR
28-SEP-1999; 99US-0156458P. PR
29-SEP-1999; 99US-0156596P. PR
04-OCT-1999; 99US-0157117P. PR
05-OCT-1999; 99US-0157753P. PR
06-OCT-1999; 99US-0157865P. PR
07-OCT-1999; 99US-0158029P. PR
08-OCT-1999; 99US-0158232P. PR
12-OCT-1999; 99US-0158369P. PR
13-OCT-1999; 99US-0159293P. PR
13-OCT-1999; 99US-0159294P. PR
13-OCT-1999; 99US-0159295P. PR
14-OCT-1999; 99US-0159329P. PR
14-OCT-1999; 99US-0159330P. PR
14-OCT-1999; 99US-0159331P. PR
14-OCT-1999; 99US-0159637P. PR
14-OCT-1999; 99US-0159638P. PR
18-OCT-1999; 99US-0159584P. PR
21-OCT-1999; 99US-0160741P. PR
21-OCT-1999; 99US-0160767P. PR
21-OCT-1999; 99US-0160768P. PR
21-OCT-1999; 99US-0160770P. PR
21-OCT-1999; 99US-0160814P. PR
21-OCT-1999; 99US-0160815P. PR
22-OCT-1999; 99US-0160980P. PR
22-OCT-1999; 99US-0160981P. PR
22-OCT-1999; 99US-0160989P. PR
25-OCT-1999; 99US-0161404P. PR
25-OCT-1999; 99US-0161405P. PR
25-OCT-1999; 99US-0161408P. PR
26-OCT-1999; 99US-0161359P. PR
26-OCT-1999; 99US-0161360P. PR
26-OCT-1999; 99US-0161361P. PR
28-OCT-1999; 99US-0161920P. PR
28-OCT-1999; 99US-0161922P. PR
28-OCT-1999; 99US-0161993P. PR
29-OCT-1999; 99US-0162142P. PR

Alignment Scores:
Pred. No.: 9,39e-74
Score: 923.00
Percent Similarity: 65.58%
Best Local Similarity: 42.09%
Query Match: 33.21%
DB: 3

Length: 1685
Matches: 181
Conservative: 101
Mismatch: 141
Indels: 7
Gaps: 3

US-10-051-909-36 (1-553) x AAC45298 (1-1685)

QY 100 PheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPhe 119
Db 207 TAATCTCAGCACATTCGTTCCGGTCTGTGTTCTTTGGTTCGGTCTTGTGCGGGATAC 266
QY 120 SerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSer 139
Db 267 TCTTCACCTGCTCAGCTGCAATAGGAATGACCTTTTCATTGACTATAGCTAGTGTTC 326
QY 140 AlaPheGlySerLeuSerAsnValGlyMetValGlyAlaIleAlaSerGlyGlnMet 159
Db 327 CTCCTTTGGTCTTTACTAATTTTGGCCCAATGATCGGTGCTATACAGAGCGGCTATA 386
QY 160 AlaGluTyrIleGlyValGlySerLeuMetIleAlaIleProAsnIleGly 179
Db 387 GCTGATTTAGTTGGAAGAAAGCGGCGATGAGATTCCTCTGCAATTTGTGTAGTCGG 446
QY 180 TrpLeuAlaIleSerPheAlaIlyAspAlaSerPheLeuTyrMetGlyArgLeuGlu 199
Db 447 TGGCTAGCAATCATCTTTGCCAAGGAGTAGTGGCTCTGGACCTTGGAGAGTGGCAACG 506
QY 200 GlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPro 219
Db 507 GGATATGNAUGGGAGCATTTCTATGTGGCCAACTTTATAGCAGAAATTCACCT 566
QY 220 GlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeu---Ser-ValThrPheGlyI 238
Db 567 AAAAAGCTTTTCAGAGGGGCTCTAACCACTGAACCCAGACAGATTCGATCTGCATGGAGT 626
QY 238 ePheLeuAlaTyrLeuLeuGlyMetPheIleProIleArgLeuLeuAlaIleGlyVal 258
Db 627 GTCCGTTTCCTTCATCATTAGGCACACTAGTACGTGGAGAGCTCTTGGCAATTAATAGGA 686
QY 258 aLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTrpLe 278

XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
KW XX
OS Arabidopsis thaliana.
XX Arabidopsis thaliana.
PN WO200216655-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US025685.
XX
PR 24-AUG-2000; 2000US-0227866P.
XX 26-JAN-2001; 2001US-0264647P.
PR 22-JUN-2001; 2001US-0300111P.
XX
PA (SCRI) SCRIPTS RES INST.
PPA (SYGN) SYNGENTA PARTICIPATIONS AG.
PI Harper JF, Krebs J, Wang X, Zhu T;
XX WPI; 2002-304127/34.
DR
XX Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.
XX
PS Claim 144; SEQ ID NO 2331; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
SQ Sequence 1473 BP; 372 A; 292 C; 352 G; 457 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.3e-73 Length: 1473
Score: 920.50 Matches: 186
Percent Similarity: 61.69% Conservative: 91
Best Local Similarity: 41.43% Mismatches: 159
Query Match: 33.12% Indels: 13
DB: 6 Gaps: 3

US-10-051-909-36 (1-553) x ABZ14526 (1-1473)

QY 98 SerAlapHeLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSer 117
||| ||| ||| :
Db 151 TCTGTTCCTCAGACGCATTTGGTCCGATCATCGGCTCTCTGTGTACC GGTTGTGCCGTT 210

QY 118 GlyPheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGlu 137
||| ||| ||| :
Db 211 GGTTCCTTCATCGGTGCACAAGCAGGGATTACCAAGATTATTACTCTCCGTTCGAGAA 270

QY 138 PheSerAlapHeGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGly 157
||| ||| ||| :
Db 271 TACTCAATGTTCCGGTCGATCTTCACATTAGGAGGCTTGATCGGTGCAGTATTTCAGCGGT 330

QY 158 GluMetAlaGluTyrlleGlyArgGlyGlySerLeuMetIleAlaAlaIleProAsnIle 177
||| ||| ||| :
Db 331 AAAGTCGTGTATGTTCTGGGAAGAACAACGGACGATGTGTTTTCGGAATTCTTCTGTATC 390

QY 178 IleGlyTtpLeuAlaIleSerPheAlaIleAspAlaSerPheLeuTyrMetGlyArgLeu 197
||| ||| ||| :
Db 391 ACAGCTGGCTTTCGTAGCATTCGCTCAGAAATCAATGTGGTGGCTGTGGAAGATTG 450

QY 198 LeuGluGlyPheGlyValGlylleIleSerTyrThrValProValTyrIleAlaGluIle 217
||| ||| ||| :

Alignment Scores:	4.32e-73	Length:	1755
Pred. No.:		Matches:	191
Score:	916.00	Conservative:	96
Percent Similarity:	61.95%	Mismatches:	165
Best Local Similarity:	1.25%	Indels:	12
Query Match:	32.96%	Gaps:	3
DB:	3		

Db	1368	TTCAACTTCTCATGAGCTGGAGCTCTCCAGGTACATTCTATTGTGTTACTCGGCTTTTGCA	1427
Qy	512	AlaPheThrIleValPheValValLeuTrpValProGluThrllysGly---***AsnSerA	531
Db	1428	GCTGCGACGATATATTTTGTGGCGAAGATGTCCTCCAGACGAAAGGGAAGACACTGGAA	1487
Qy	531	rgGlyAspThrIlePheValSerLeuSerIleGlnArgGlnLeuGln---TrpLeuProG	550
Db	1488	GAGATCCAGCTTGTATTTCGAAGAGAACATAGAAGAAAGATACATCAATGTTATGGG	1547
Qy	550	lucYs 551	
Db	1548	AATGT 1552	
RESULT 11			
ID	ABZ13822	standard; DNA; 1359 BP.	
XX	AC	ABZ13822;	
XX	DT	21-JAN-2003 (first entry)	
XX	DE	Arabidopsis thaliana stress regulated gene SEQ ID NO 1627.	
XX	KW	Arabidopsis thaliana; plant; gene; stress; transgenic; ds.	
XX	OS	Arabidopsis thaliana.	
XX	PN	WO200216655-A2.	
XX	PD	28-FEB-2002.	
XX	PF	24-AUG-2001; 2001WO-US026685.	
XX	PR	24-AUG-2000; 2000US-0227866P.	
XX	PR	26-JAN-2001; 2001US-0264647P.	
XX	PR	22-JUN-2001; 2001US-030011P.	
PA	(SCRI)	SCRIPPS RES INST.	
PA	(SYGN)	SYNGENTA PARTICIPATIONS AG.	
PI	Harper JF,	Kreps J, Wang X, Zhu T;	
PI	WPI;	2002-304127/34.	
DR	XX		
XX	PT	Identifying a stress condition to which a plant cell has been exposed and	
PT	PT	producing plants with increased tolerance to these abiotic stresses.	
XX	PS	Claim 144; SEQ ID NO 1627; 577pp + Sequence Listing; English.	
XX	CC	The invention relates to identifying a stress condition to which a plant	
CC	CC	cell has been exposed, comprising: (a) contacting nucleic acid	
CC	CC	representative of expressed polynucleotides in the plant cell with an	
CC	CC	array or probes representative of the plant cell genome; and (b)	
CC	CC	detecting a profile of expressed polynucleotides in the plant cell	
CC	CC	characteristic of a stress response. The method is useful in the	
CC	CC	production of transgenic plants, cells and seeds and in producing plants	
CC	CC	with increased tolerance to abiotic stress. The present sequence is that	
CC	CC	of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used	
CC	CC	in methods of the invention. Note: The sequence data for this patent is	
CC	CC	not represented in the printed specification but is based on sequence	
CC	CC	information supplied to Derwent by the European Patent Office	
XX	XX		
XX	XX	Sequence 1359 BP; 340 A; 238 C; 327 G; 454 T; 0 U; 0 Other;	
Alignment Scores:			
Pred. No.:	1.04e-71	Length:	1359
Score:	899.00	Matches:	188
Percent Similarity:	60.17%	Conservative:	99
Best Local Similarity:	39.41%	Mismatches:	149
Query Match:	32.95%	Indels:	42
DB:	6	Gaps:	5

US-10-051-909-36 (1-553) x ABZ13822 (1-1359)	
Qy	79 GlnSerSerPheAlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSer 99
Db	432 TyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPhe 451
Db	1060 -----GTGTACACCGGATCATTTTCATTA 1083
Qy	99 Ala-----PheLeuCysThrLeuIleValAlaLeuGlyProIle 111
Db	73 TCGCGGACATAACMACTACTCTCTCTCCTACACATTTTGTAGCCGTGCTGCTCTTT 132
Qy	112 GlnPheGlyPheThrSerGlyPheSerProThrGlnAspAlaMetValArgAspLeu 131
Db	133 GTCTTTGGCTCAGCATATAGGATATTCACCTGTTCCAGTCGATTAACCAAGGAGTTG 192
Qy	132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal 151
Db	193 AATCTCTCTGTGGCAGAGTACTGCTCTCTCGGTTCGATTCGCAATTCGAGCAATGAT 252
Qy	152 GlyAlaIleAlaSerGlyGlnMetAlaGluThrIleGlyArgGlySerLeuMetIle 171
Db	253 GCGCAGCTATGAGTGACGACGATCGCGATATGATAGCAGACGACTACTATGGGTTTC 312
Qy	172 AlaAlaIleProAsnIleIleGlyThrLeuAlaIleSerPheAlaIleSerPhe 191
Db	313 TCGGAGATGTTTTCATTTCTCGGCTGCTTACCAATCTACCTTTCTAAGGTTCGCAATTTGG 372
Qy	192 LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro 211
Db	373 CTTCAGCTGGGAGATCTTGGTGGTATGATGAGGCGGTTTTCGTTTGGTTCCT 432
Qy	212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231
Db	433 GTATACATAGCTGAATCACTCTTAAGGTCCTTCGTCGGTGGTTTACCACTGTTTCATCAG 492
Qy	232 LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrPac 251
Db	493 TTAGTATATGTTGGTGGTCTCAGTCACATATCTCTTGGGCTCTTTTAGGCTGGCGA 552
Qy	252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIle 271
Db	553 ATTTAGCTTTGATCGGAATGATACCATGATGTTGTACAAATGATGGGATTTTGTGCATC 612
Qy	272 ProGluSerProArgThrLeuAlaIleMetAsnLeuThrGluAspCysGluThrSerLeu 291
Db	613 CCAGAGTCTCTAGATGGTGGCAAAAGTTGGCAAAATGGGAAGAATTTGAGATTGCATTG 672
Qy	292 GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAla 311
Db	673 CAACGATTAGTGTGTAATCGCGAGATATTTCAATGAGTCTAATGAGATCAAGATTAT 732
Qy	312 ValAlaSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyr 331
Db	733 ACGCGAAGACTTACAGATTTATCAGAAGGTAGTATAGTATCTGTTCCAGCCACAGTAT 792
Qy	332 ArgThrProLeuLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsn 351
Db	793 GCTAAATCTCTAGTGTGGGAGTGGTCTGATGGTGTTCGACAAATTTGAGGGGGTTAAT 852
Qy	352 GlyValLeuPheTyrAlaSerSerIlePheIleAlaAlaGlyValThrAsnSerAspLeu 371
Db	853 GGAATTGCTTTTACGCTAGCTCTATTTTCGAATCTGCTGGGTT--TCTTCCAAAT 909
Qy	372 AlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeu 391
Db	910 GGAATGATGCGTGGTGGTGCATACATCCCTATGACAACTTTAGTGTTACTGTTGATG 969
Qy	392 AspArgAlaGlyArgGlyIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeu 411
Db	970 GATAAATCAGGAGACGACCATTAATTTCTCTACTGGAACATGATCGGGTGT 1029
Qy	412 LeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyr 431

1030 TTCCTCGTGGTTTATCATTTTCATTACAG----- 1059

Qy 432 TyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPhe 451

Db 1060 -----GTGTACACCGGATCATTTTCATTA 1083

Qy 452 GlyMetGlyAlaIleProThrLeuMetSerGluIleLeuProValSerIleLysSer 471

Db 1084 GGGATGGTGGGATTCCTTGGTCTATATGTCAGAGATATTTCCATAGACATAAAGGA 1143

Qy 472 LeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThr 491

Db 1144 TCAGCCGGAAGCTCGTGACTGTTTGTAGCTGGGTGGATCATGATGATTATATCTTTTACA 1203

Qy 492 ThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValValSer 511

Db 1204 TTTAATCTTCTTAATGAATGGATCCGACGAGACGTTTATGTTTCGCTACTCTTTGT 1263

Qy 512 AlaPheThrIleValPheValValLeuTrpValProGluThrLysGly***AsnSerArg 531

Db 1264 GGGGCTACTGTTATTTTGTAGCAAACTCGTACCAGAAACCAAGGTCG-CACACTTGA 1322

Qy 532 GlyAspThrIlePheValSerLeuSerIleGlnArgGlnLeuGlnTrpLeu 548

Db 1323 GGAATCCCAATATTC-----GATTGGTTA 1346

RESULT 12

ADM47939

ID ADM47939 standard; DNA; 1768 BP.

XX ADM47939;

AC ADM47939;

XX 03-JUN-2004 (first entry)

XX Polynucleotide sequence #357 useful in producing transgenic plants.

DE Plant; maize; rice; sorghum; thale cress; soybean; wheat; transgenic;

KW osmotic stress; sugar transport; cell cycle pathway; plant height;

KW carbohydrate transport; crop productivity; plant growth;

KW stress resistance; disease resistance; insect resistance; heat tolerance;

KW nitrogen assimilation; water stress tolerance;

KW photosynthetic carbon fixation; virus resistance; gene therapy; gene; ds.

XX Glycine max.

OS US2003233670-A1.

XX 18-DEC-2003.

XX 04-DEC-2002; 2002US-00310154.

XX 04-DEC-2001; 2001US-0337358P.

XX (EDGE/) EDGERTON M D.

XX (CHOM/) CHOMET P S.

XX (LACC/) LACCETTI L B.

XX Edgerton MD, Chomet PS, Laccetti LB;

XX WPI; 2004-061374/06.

XX P-PSDB; ADM48307.

XX New polynucleotide, useful for manipulating plant protein quality,

PT improving plant growth, yield and crop productivity or grain composition

PT or producing plants with improved properties.

XX Claim 1; SEQ ID NO 357; 144pp; English.

XX The present invention relates to polynucleotide sequences, and the

CC proteins they encode. The sequences are isolated from a variety of

CC organisms such as plants (e.g. maize, rice, sorghum, thale cress,

CC soybean, and wheat), cyanobacteria, bacteria, yeast and other fungi. The

CC polynucleotide and polypeptide sequences of the invention are useful in

the production of transgenic plants that have improved properties. Also disclosed are methods of producing fertile transgenic plants, preferably maize, with desired phenotypes. The polynucleotide and polypeptide sequences are useful for improving plants by providing protection against osmotic stress, improving altering sugar transport and/or metabolism, modifying the cell cycle pathway, reducing plant height, modifying carbohydrate transport, improving crop productivity, improving plant growth and stress resistance, improving disease resistance, improving insect resistance, improving cold or heat tolerance, improving nitrogen assimilation, improving stalk strength, improving water stress tolerance, improving photosynthetic carbon fixation, improving biotic and abiotic stress resistance, improving resistance to oxidative stress, providing increased vigour, reducing senescence, and conferring virus resistance. The present sequence represents a polynucleotide sequence of the invention. Note: The sequence data for this patent is not provided in the printed specification but is obtained in electronic format from the USPTO website at seqdata.uspto.gov.

Sequence 1768 BP; 516 A; 300 C; 384 G; 568 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5.4e-71 Length: 1768
Score: 893.00 Matches: 191
Percent Similarity: 56.25% Conservative: 97
Best Local Similarity: 37.30% Mismatches: 173
Query Match: 32.13% Indels: 51
DB: 12 Gaps: 4

US-10-051-909-36 (1-553) x ADM47939 (1-1768)

Qy 43 SerAsnArgGly-GlyAlaGlyAlaGlyGluSerGlySerAspHisaspGlyValle 62
Db 119 TCAACAGAGAGTCTGCTGCTGCTTGTAAAGAAATGATCAGAT-----164
Qy 62 uArgArgProLeuLeuAsnThrGlySerTyrArgMetSerSerArgGlnSerSerPh 82
Db 165 -----AAGAGTGT 172
Qy 82 eAlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCy 102
Db 173 GAAAAATGGATCCATTCGATGGTGTG-----CTTAG 205
Qy 102 sThrLeuLeuValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerPr 122
Db 206 CACACTTGTGCTGCTGCTGCTTTCACATTTGGAACTTGTGTGGGCTATTTCAGCAC 265
Qy 122 oThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPheGll 142
Db 266 CACTCAGCGGCTATCAGGCAGATCTTAATCTCTCTGCTGAGTTTCCATGTTGG 325
Qy 142 ySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTy 162
Db 326 TTCATTAGTACCATTCGTGCAATGCTTGGAGCTATAACAGCGGCAGGATTACAGATT 385
Qy 162 rIleGlyArgGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAl 192
Db 386 CATTGGCCGCAAGGCAATGAGATTTCAACAGGATTTTGCATTACAGATGGATGAGC 445
Qy 182 aileSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGlyPheGl 202
Db 446 TGTCTTCTTCTAAGGTTCTTACTCCCTGACTTCGGAAGATTTTTCACAGGTTACGG 505
Qy 202 yValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMe 222
Db 506 CATTGGAGTATTCATCAGCGGTTCTGTATATATAGCAGAAATAGCACCACCAATCT 565
Qy 222 tArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlar 242
Db 566 TCGAGGAGGACTAGCAACAAATCAGCTTTGATTGTTACTGGAGGGTCAGTCTCAT 625
Qy 242 rLeuLeuGlyMetPheIleProThrArgLeuLeuAlaValIleGlyAlaLeuProCysTh 262
Db 626 CTTATTAGGAGTGTATAAATTGGAGAGAACTTGCATTAGCGGGGCTAGTCTGCTGCAT 695

Qy 262 rMetLeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeuAlaLeuMetAs 282
Db 686 TTGCTTGTCTGTTGTTGTTGTTGTTATCCCTCAGTCCCCCAGATGGTGGTAAAGTTGG 745
Qy 282 nLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleTh 302
Db 746 CCGTGAAGAAAGAAATTTCAACTAGCTTTAAAGTAGACTTCGGGGTAAAGATGCTGATATTC 805
Qy 302 rThrGluValAsnAspIleLysArgAlaValAlaSerSerLysArgThrThrIleSe 322
Db 806 TGATGAAGCTGCTGAAATTCGGATTATTTGAAACCTCTTCAAGTCTTCTTAAGACTAA 865
Qy 322 rPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGlyLeuLe 342
Db 866 GCTGTGGAATTTGTTCCAAAGCAATATGTCACCTGCTAGTATTGGGTGCGCTTAAT 925
Qy 342 uValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLy 362
Db 926 GGCATGTCAACAATCTGTTGGAATTAATGGCATAGGATTTACACAGCTGAGATTTTGT 985
Qy 362 eAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValle 382
Db 986 AGCAGCTGAGCTTTCTTCAGAAAGCTGTTACATAGCATATGCTGTATACAGATTCC 1045
Qy 382 uAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIle 402
Db 1046 ATTTACATTATTGGGAGCCATTTTGATGCAAGTCTGGAAGAAGACCTCTTGTAAATGCT 1105
Qy 402 sSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeuLysAs 422
Db 1106 TTCCGACCTGGACATTTCTTAGGTGCTTGTGCTGCTTCTTCTTCTTCTTCTTCTTCT 1165
Qy 422 pAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGlyI 442
Db 1166 C-----CAAAGCTTATTGCTGAGTGGGTACCTATATTAGCATTGTGCGGT 1213
Qy 442 eValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProThrPleuMetMetSe 462
Db 1214 TGTGATCTACATAGCAGCATTTTCAATTTGGCTCGGATCAGTTCCATGGGTGATAATGTC 1273
Qy 462 rGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTr 482
Db 1274 TGAGATATTCCCATACATTTGAAGGGGACTGCTGGAAGCTTGTGTTTGGTGGCTTG 1333
Qy 482 pLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGly 502
Db 1334 GCTAGGAGCTGGGTAGTTTTCATATACCTTTTCAACTTTCTTATGAGCTGGAGTTCTCCTGG 1393
Qy 502 yThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValLeuTpvVa 522
Db 1394 TACATTGTTTGTATGCTGGATGTTTCTCTTAACTATTCTATTGTTAGCAAAATAGT 1453
Qy 522 lProGluThrLysGly****AsnSerArgGlyAspThrIlePheValSerLeuSerIleGl 542
Db 1454 CCCAGAAACCAAGGA-----1469
Qy 542 nArgGlnLeuGlnTrpLeuProGluCysLeuSer 553
Db 1470 -AAACCTTTGGAAGAGATCCAGGCTTGATTAGT 1502
RESULT 13
AB213486
ID AB213486 standard; DNA; 1407 BP.
XX AB213486;
AC AB213486;
XX 21-JAN-2003 (first entry)
XX Arabidopsis thaliana stress regulated gene SEQ ID NO 1291.
DE Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX

OS Arabidopsis thaliana.
XX W0200216655-A2.
XX 28-FEB-2002.
XX 24-AUG-2001; 2001WO-US026685.
XX 24-AUG-2000; 2000US-0227866P.
XX 26-JUN-2001; 2001US-0264647P.
XX 22-JUN-2001; 2001US-0300111P.
XX (SRI) SCRIPPS RES INST.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX Harper JF, Krops J, Wang X, Zhu T;
XX WPI; 2002-304127/34.
XX Identifying a stress condition to which a plant cell has been exposed and
XX producing plants with increased tolerance to these abiotic stresses.
XX Claim 144; SEQ ID NO 1291; 577pp + Sequence Listing; English.
XX The invention relates to identifying a stress condition to which a plant
XX cell has been exposed, comprising: (a) contacting nucleic acid
XX representative of expressed polynucleotides in the plant cell with an
XX array or probes representative of the plant cell genome; and (b)
XX detecting a profile of expressed polynucleotides in the plant cell
XX characteristic of a stress response. The method is useful in the
XX production of transgenic plants, cells and seeds and in producing plants
XX with increased tolerance to abiotic stress. The present sequence is that
XX of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
XX in methods of the invention. Note: The sequence data for this patent is
XX not represented in the printed specification but is based on sequence
XX information supplied to Derwent by the European Patent Office
XX
XX Sequence 1407 BP; 350 A; 261 C; 331 G; 465 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 9,92e-70 Length: 1407
Score: 877,50 Matches: 179
Percent Similarity: 59.06% Conservative: 98
Best Local Similarity: 38.17% Mismatches: 166
Query Match: 31.58% Indels: 26
DB: 6 Gaps: 5
US-10-051-909-36 (1-553) x AB213486 (1-1407)
QY 66 LeuLeuAsnThrGlySerTyrArg-MetSerSerArgGlnSerSerPheAlaProG1 85
Db 12 CTCAGAAACTTGGAGCTGGTTTGCTACTGAGGAAAGAACCAACGAC----- 60
QY 85 yThrSerSerMetAlaValLeuArgGluSerHisValSerAla-----PheLeuCysTh 103
Db 61 -----ATCAACGATGCTGATATCAGTCTGTGTACTTTTCAGTAC 101
QY 103 rLeuLeuValAlaLeuGlyProLeuGlnPheGlyPheThrSerGlyPheSerSerProTh 123
Db 102 TTTGGTCTCTGTGTGGCTCTTCTGCTTCGGTGTGGCGCAGGTTATTTCATCAGTTGC 161
QY 123 rGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySe 143
Db 162 TCAACAGGATCATAAATGATTAGGTCCTCTGTGTGCAATATCTCCATGTTTGGTTC 221
QY 143 rLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGlyTril 163
Db 222 AATCATGACTTTTGGAGGAATGATTGGTGCCATCTTCAGGCGGAAAGTTGCAGATCTCAT 281
QY 163 eGlyArgGlySerLeuMetIleAlaIleAlaProAsnIleIleGlyTrpLeuAlar1 183
Db 282 GGSTCGAAAGGGGACTATGTTGGTTGCTCAAAATTTCTGCACTCTCGTGGTGGTTCAGT 341

RESULT 14

Qy	499	rValGlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValVa	519	PR	16-JUN-1999;	99US-0139452P;
		::: :::		PR	16-JUN-1999;	99US-0139453P;
Db	1275	CGCATCAGGAACCTCTGTATCTTTACTATATGTTGGCGGTATAGCTTTATTTA	1334	PR	17-JUN-1999;	99US-0139492P;
		::: :::		PR	18-JUN-1999;	99US-0139454P;
Qy	519	lLeutrpValProGluThrLysGly	527	PR	18-JUN-1999;	99US-0139455P;
		::: :::		PR	18-JUN-1999;	99US-0139456P;
Db	1335	TGCGATGGTACCGAACTAAAGGA	1359	PR	18-JUN-1999;	99US-0139457P;
		::: :::		PR	18-JUN-1999;	99US-0139458P;
RESULT 15				PR	18-JUN-1999;	99US-0139459P;
AAC46009				PR	18-JUN-1999;	99US-0139460P;
ID	AAC46009	standard; DNA; 1642 BP.		PR	18-JUN-1999;	99US-0139461P;
XX				PR	18-JUN-1999;	99US-0139462P;
AC	AAC46009;			PR	18-JUN-1999;	99US-0139463P;
XX				PR	18-JUN-1999;	99US-0139750P;
DT	18-OCT-2000	(first entry)		PR	18-JUN-1999;	99US-0139763P;
XX				PR	21-JUN-1999;	99US-0139817P;
DE	Arabidopsis thaliana	DNA fragment SEQ ID NO: 48579.		PR	22-JUN-1999;	99US-0139899P;
XX				PR	23-JUN-1999;	99US-0140353P;
KW	Hybridisation assay; genetic mapping; gene expression control;			PR	23-JUN-1999;	99US-0140354P;
KW	protein identification; signal transduction pathway; metabolic pathway;			PR	24-JUN-1999;	99US-0140695P;
KW	promoter; termination sequence; ss.			PR	28-JUN-1999;	99US-0140682P;
XX				PR	29-JUN-1999;	99US-0140991P;
OS	Arabidopsis thaliana.			PR	30-JUN-1999;	99US-0141287P;
XX				PR	01-JUL-1999;	99US-0141842P;
PN	EP1033405-A2.			PR	01-JUL-1999;	99US-0142154P;
XX				PR	02-JUL-1999;	99US-0142055P;
PD	06-SEP-2000.			PR	06-JUL-1999;	99US-0142390P;
XX				PR	08-JUL-1999;	99US-0142803P;
Pf	25-FEB-2000; 2000EP-00301439.			PR	09-JUL-1999;	99US-0143202P;
XX				PR	12-JUL-1999;	99US-0143777P;
XX	25-FEB-1999;	99US-0121825P.		PR	13-JUL-1999;	99US-0143542P;
PR	05-MAR-1999;	99US-0123180P.		PR	14-JUL-1999;	99US-0143624P;
PR	09-MAR-1999;	99US-01233548P.		PR	15-JUL-1999;	99US-0144005P;
PR	23-MAR-1999;	99US-0125788P.		PR	16-JUL-1999;	99US-0144085P;
PR	25-MAR-1999;	99US-0126264P.		PR	16-JUL-1999;	99US-0144086P;
PR	23-MAR-1999;	99US-0126785P.		PR	19-JUL-1999;	99US-0144325P;
PR	01-APR-1999;	99US-0127462P.		PR	19-JUL-1999;	99US-0144331P;
PR	08-APR-1999;	99US-0128234P.		PR	19-JUL-1999;	99US-0144332P;
PR	16-APR-1999;	99US-0128714P.		PR	19-JUL-1999;	99US-0144333P;
PR	19-APR-1999;	99US-0130077P.		PR	19-JUL-1999;	99US-0144334P;
PR	21-APR-1999;	99US-0130049P.		PR	20-JUL-1999;	99US-0144335P;
PR	23-APR-1999;	99US-0130510P.		PR	20-JUL-1999;	99US-0144632P;
PR	23-APR-1999;	99US-0130891P.		PR	20-JUL-1999;	99US-0144684P;
PR	30-APR-1999;	99US-0132048P.		PR	21-JUL-1999;	99US-0145086P;
PR	04-MAY-1999;	99US-0132484P.		PR	21-JUL-1999;	99US-0145088P;
PR	05-MAY-1999;	99US-0132485P.		PR	22-JUL-1999;	99US-0145085P;
PR	06-MAY-1999;	99US-0132486P.		PR	22-JUL-1999;	99US-0145087P;
PR	07-MAY-1999;	99US-0132487P.		PR	22-JUL-1999;	99US-0145089P;
PR	11-MAY-1999;	99US-0132863P.		PR	23-JUL-1999;	99US-0145192P;
PR	14-MAY-1999;	99US-0134218P.		PR	23-JUL-1999;	99US-0145224P.
PR	14-MAY-1999;	99US-0134219P.		PR	26-JUL-1999;	99US-0145276P;
PR	14-MAY-1999;	99US-0134221P.		PR	27-JUL-1999;	99US-0145913P;
PR	14-MAY-1999;	99US-0134370P.		PR	27-JUL-1999;	99US-0145918P;
PR	18-MAY-1999;					


```
QY 423 nileSerGlnAspSerAsnSerTyrTyrIleLeu-----ThrMetIleSerLe 439
Db 1060 -----TCATAGGCATGCTTGATGAGCTCACTCCGATTTTCACATG 1100
QY 439 uValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTyrLe 459
Db 1101 TATCGGTGTAGTGGGTTTCATCTCTTCATTGCGGTAGGCATGGGAGGCTTACCATGGAT 1160
QY 459 uMetMetSerGluIleLeuProValSerIleLyssSerLeuGlyGlySerIleAlaThrLe 479
Db 1161 CATCATGCTCGAGATATTCCCAATGATGTTAAAGTTTCTGCTGGGACTCTGGTTACCTT 1220
QY 479 uAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSe 499
Db 1221 AGCCAACTGCTCCTTTGGTGGATTGTTGCTTTCGGCTACAACTTCATGCTAGAGTGGAA 1280
QY 499 rValGlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValVa 519
Db 1281 CGCATCAGGAACGTTCTTGATCTCTTTACTATATGTGTGGCGGTATAGTCTTTATTTA 1340
QY 519 lLeuTrpValProGluThrLySGly 527
Db 1341 TGGCATGGTACCGAAGAACTAAAGGA 1365
```

Search completed: October 13, 2004, 12:24:30
Job time : 717.434 secs

is Page Blank (uspto)

QY	86	ThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCys---ThrLeu	104
DB	30	ACTAGTACAATGACA-----GAGATAAGGTCACTGGGACCCTGTTTCTCACTGTC	80
QY	105	IleValAla---LeuGlyProIleGlnPheGlyPheThrSerGlyPheSerProThr	123
DB	81	ATCACTGCTGCTGGTTCCTTCAGTTTGGATATGACATTGGTGTGATCAATGCACCT	140
QY	124	GlnAspAlaMetValArgAspLeu-----	131
DB	141	CAACAGGTAATAATATCTCACTATAGACATGTTTGGTGTTCCTCACTGGATGACCGAAA	200
QY	131	-----	131
DB	201	GCTATCAACAACACTATGTTATCAACAGTACAGATGAAGTCCCAACAAATCTCACTCAATG	260
QY	131	-----	131
DB	261	AAACCAAAACCAACCCCTTGGCTGGAGAGAGACTGTGGCAGCTGCTCAACTAATCACC	320
QY	132	-----AsnLeuSerIleSerGlnPheSerAlaPheGlySerLeuSerAsnValGly	148
DB	321	ATGCTCTGGTCCCTGCTGTATCCAGCTTTTGCA-----GTTGGT	359
QY	149	GlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrlleGlyArgLysGlySer	168
DB	360	GGAATGACTGCATCATCTTTTGGTGGTGGCTTTGGGACACACTTGGAGAATCAAGACC	419
QY	169	LeuMetIleAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLys---	187
DB	420	ATGTTAGTAGCAAAACATTCTGCTCATATAGTTGGAGCTCTCTGATGGGTTTTCAAAATG	479
QY	188	-----AspAlaSerPheLeuTyrrMetGlyArgLeuGluGlyPheGlyValGlyIle	205
DB	480	GGACCATCTCATATATATAATTGCTGGAGAAGCATATCAGGACTATATTGTGGGCTA	539
QY	206	IleSerTyrrValProValTyrlleAlaGluIleSerProGlnAsnMetArgGlyAla	225
DB	540	ATTTCAGCCCTGGTTCCTATGATATCGGTGAATTTGCTCCAAACGCTCTCAGGGGAGCA	599
QY	226	LeuGlySerValAsnGlnLeuSerValThrPheGlyIlePhe-----	239
DB	600	CTTGGCACTTTTCATCAGCTGGCCATCGTCACGGCATTCTTATTAGTCAGATTTATGGT	659
QY	240	LeuAlaTyrrLeuLeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeu	259
DB	660	CTTGAATTATCTTTGGCAATTATGATCTGTGGCACATCTTGCTGGCTGCTGTGTGTG	719
QY	260	ProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeu---	278
DB	720	CGAGCCATCCTTCAGTCTCTGCTACTCTTTTCTGTCCAGAAAGCCACAGATCCTTAC	779
QY	279	AlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGlu	298
DB	780	ATCAAGTTAGTAGAGGAAGTCAAGCAAAACAAAGCTTGAAAAGACTCAGAGGATATGAT	839
QY	299	ThrAspIleThrThrGluValAsnAspIleLys-----ArgAlaValAlaSerSerSer	316
DB	840	---GATGTCACCAANGATATTATGAATGAGAAAAGAAAGAGAGATCGATGTCAG	896
QY	317	LysArgThrThrIleSerPheGlnGluLeuAsnGlnLysTyrrArgThrProLeuLeu	336
DB	897	CAGAAGGTCCTATA---ATTCACTGCTTCCCAATTCACGCTACCGCAGCCCTATTCTA	953
QY	337	LeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrr	356
DB	954	GTGGCACTGATGCTGCATGTGGCTCAGCAATTTTCCGGAATCAATGGCAATTTTTTACTAC	1013
QY	357	AlaSerSerIlePheLysAlaIleGlyValThrAsnSerAspLeuAlaThrCysSerLeu	376
DB	1014	TCACCCACGATTTTTCAGCGGTGTGATCTCAGCAAAACCTGTTTATGTCAACCATTTGAGTT	1073

Qy	377	GlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArg	396
Db	1074	GGGGCTGTAACATGGTGTTCACCTGCTGCTCTGTATTCCTTGTGGAGAAGGCAGGCGGA	1133
Qy	397	ArgIleLeuLeuIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerVal	416
Db	1134	CGTTCTCTCTTCTAAATGGAAATGAGTGGAGATTGTTTGTGCCATCTTCATGTCAGTG	1193
Qy	417	ValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMet	436
Db	1194	GGACTTGTGCTG-----CTGAATAAGTTCTCTTGGATGAGTTAT	1232
Qy	437	IleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIle	456
Db	1233	GTGAGCATGATGACCATCTTCCTCTTTGTGAGCTTCTTGAATGTGGCCAGGCCCGATC	1292
Qy	457	ProTrpLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlyLysIle	476
Db	1293	CCCTGGTTTCATGTGGCTGAGTGTTCAGTCAAGGACCACGTCCTGCTGCTTTAGCAATA	1352
Qy	477	AlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeu	496
Db	1353	GCTCGCATTCAGCAATTGGACCTGCCAATTCATTGTAGCTCTGTGTTTCCAGTACATGGG	1412
Qy	497	ThrTrpSerValGlyGlyThrPheLeuSerTyrMet-----ValValSerAlaPhe	513
Db	1413	GACTTCTGT-----GGACCTTATGTGTTTTCCTCTTCTCGGAGTGTCTCTGGCTTT	1466
Qy	514	ThrIleValPheValValLeuTrpValProGluThrLysGly	527
Db	1467	ACCCCTG---TTTACATTTTAAAGTTCCAGAAACCAAAAGGA	1505

RESULT 2

```

US-09-919-172-79
; Sequence 79, Application US/09919172
; Patent No. 6673545
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 79
; LENGTH: 5227
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673545 410785.1
; NAME/KEY: un28
; LOCATION: 4928, 4934, 4939, 4944, 4973, 4992
; OTHER INFORMATION: a, t, c, g, or other
US-09-919-172-79

```

Alignment Scores:	
Pred. No.:	1.6e-50
Score:	560.50
Percent Similarity:	49.03%
Best Local Similarity:	29.77%
Query Match:	20.17%
DB:	4
	Gaps:
	15
	Indels:
	94
	Mismatches:
	169
	Conservative:
	99
	Matches:
	153
	Length:
	5227

US-10-051-909-36 (1-553) x US-09-919-172-79 (1-5227)

Qy 86 ThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCys--ThrLeu 104
 |||||:::||| |||||:::||| |||||:::||| |||||:::|||
Db 30 ACTAGTACAATGAC-----GAAGATAAGGTCACTGGGACCCTGGTTTCACGTGC 80

QY 105 IleValAla---LeuGlyProIleGlnPheGlyPheThrSerGlyPheSerProThr 123
 Db 81 ATCACTGCTGTGGTGGTCTCTCCAGTTGGATATGACATTGGTGTGATCAATGACACT 140
 QY 124 GlnAspAlaMetValArgAspLeu----- 131
 Db 141 CAACAGGTAATAATATCTCACTNAGACATGTTTGGGTGTTCCACTGGATGACCGAAA 200
 QY 131 ----- 131
 Db 201 GCTATCAACAACTATGTTATCAACAGTACAGATGAAGTGGCCACATCTCATACTCAATG 260
 QY 131 ----- 131
 Db 261 AACCCAAAACCAACCCCTGGGTGGAGAGACACTGTGGCAGCTGCTCAACTAATACACC 320
 QY 132 -----AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGly 148
 Db 321 ATGCTGTGGTCCCTGCTGTATCCAGCTTGCA-----GTTGCT 359
 QY 149 GlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgIysGlySer 168
 Db 360 GGAATGACTCATCTATCTTTGGTGGTGGTGGGACACACTTGGAAAGATCAAGCC 419
 QY 169 LeuMetIleAlaAlaIleProAsnIleIleGlyTyrLeuAlaIleSerPheAlaLys--- 187
 Db 420 ATGTTAGTACAAACATCTGTCTATAGTTGGAGCTCTCTGATGGGTTTTCAAAATG 479
 QY 188 -----AspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIle 205
 Db 480 GGACCATCTCATATCTATAATTTGCTGGAGAGCATATCAGCATATATTGGGCTA 539
 QY 206 IleSerTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAla 225
 Db 540 ATTTCAGGCTGTGCTATGATATGATGCTGGGAAATGCTCAACCGCTCTCAGGGAGCA 599
 QY 226 LeuGlySerValAsnGlnLeuSerValThrPheGlyIlePhe----- 239
 Db 600 CTGGCACTTTTCATCAGCTGGCATCGTCACGGCATTTCTATTAGTCAGATTATGGT 659
 QY 240 LeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaValIleGlyAlaLeu 259
 Db 660 CTTGAATTTATCTGGGCAATATGATCTGTGGCACAATCTTGTGGCTGTCTGGTGTG 719
 QY 260 ProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTyrLeu--- 278
 Db 720 CGAGCCATCCTCTAGTCTGTCTACTCTTTTCTGTCAGAAAGCCCAATACCTTTAC 779
 QY 279 AlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGlu 298
 Db 780 ATCAAGTTAGTAGAGGAAGTCAAGCAAAACAAAGCTTGAAGAGACTCAGAGGATATGAT 839
 QY 299 ThrAspIleThrThrGluValAsnAspIleLys-----ArgAlaValAlaSerSerSer 316
 Db 840 ---GATGTCAACCAAGATATTAATGAATGAGAAAGAAAGAGAGAGAGATCGAGTGAG 896
 QY 317 LysArgThrThrIleSerPheGlnGluLeuAsnGlnLysIysTyrArgThrProLeuLeu 336
 Db 897 CAGAAAGTCTCTATA---ATTGACTCTTCCCAATTTCCAGTACCGACAGCCTATTCTA 953
 QY 337 LeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyr 356
 Db 954 GTGGCACTGATGCTGCTGTGGCTCAGCAATTTTCCGGAATCAATGTCATTTTACTAC 1013
 QY 357 AlaSerSerIlePheLeuAlaIleGlyValThrAsnSerAspLeuAlaThrCysSerLeu 376
 Db 1014 TCACACGACATTTTTCAGCGCTGGTATCAGCAAAACCTGTTTATGCAACCATTTGGATT 1073
 QY 377 GlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArg 396
 Db 1074 GGCCTGTAAACATGGTTTTCACCTGCTGTCTCTGATTTCTTGTGGAGAGGAGGCGCA 1133
 QY 397 ArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerVal 416

Db 1134 CGTTCTCTCTTTAAATGGGAATGAGTGGGATGTTTGTTCGCACTTTTCATGTCAGTG 1193
 QY 417 ValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMet 436
 Db 1194 GGACTGTGCTGCTGAATAAGTTCTCTTG-GAC-----TGAGTTACT 1234
 QY 437 IleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIle 456
 Db 1235 GTGACATGATAGCACTCTCTCTTTGTGACGTTCTTTGAAATGGGCCAGGCCGATC 1294
 QY 457 ProTrpLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIle 476
 Db 1295 CCCTGGTTCATGGTGGCTGAGTTTTCAGTCAAGGACACCGCTCTGCTGCTTAGCAATA 1354
 QY 477 AlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeu 496
 Db 1355 GCTGATTCAGCAATTCGACCTGCAATTCATTGTTAGTCTGTGTTCCAGTACATTGCG 1414
 QY 497 ThrTrpSerValGlyGlyThrPheLeuSerTyrMet-----ValValSerAlaPhe 513
 Db 1415 GACTTCTGT-----GGACCTTATGTGTTTCTCTTCTGCTGGAGTCTCTGCGCTTT 1468
 QY 514 ThrIleValPheValValLeuTrpValProGluThrLysGly 527
 Db 1469 ACCCTG---TTTACATTTTAAAGTTCCAGAACCAAGGA 1507
 RESULT 3
 US-09-291-922-25
 ; Sequence 25, Application US/09291922
 ; Patent No. 6383776
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Hitz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Kinney, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB-1163
 ; CURRENT APPLICATION NUMBER: US/09/291.922
 ; CURRENT FILING DATE: 1999-04-14
 ; EARLIER APPLICATION NUMBER: 60/083.044
 ; EARLIER FILING DATE: April 24, 1998
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 25
 ; TYPE: DNA
 ; ORGANISM: Triticum aestivum
 ; US-09-291-922-25
 Alignment Scores:
 Pred. No.: 6,87e-51 Length: 2089
 Score: 557.50 Matches: 135
 Percent Similarity: 49.73% Conservative: 97
 Best Local Similarity: 28.97% Mismatches: 187
 Query Match: 20.08% Indels: 47
 DB: 3 Gaps: 7
 US-10-051-909-36 (1-553) x US-09-291-922-25 (1-2089)
 QY 99 AlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGly 118
 Db 137 GCCTTCGCTCGCCATCTCCCTCCATGACCTCCTCCTCCTCGCTACGATCGCG 196
 QY 119 PheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPhe 138
 Db 197 GTGATGAGCGGAGCTGCTGTACATCCAGAGGATCTGAAGATCAACGACACCCAGCTG 256
 QY 139 SerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGln 158
 Db 257 GAGGTCTCATGGCATCTCAACGCTACTCGCTCATTTGGCTCTCTCGCGCGGGCGG 316
 QY 159 MetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIle 178

```

Db      317  ACCTCCAGTGGATCGCGCGGCTTACCATCGTCTTCGCCGCGCATCTTCTTCGCG 376
Qy      GlyThrPheLeuSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeu 198
Db      377  GCGCCCTCATCATGCGGCTTCTCGCTCACTACGCGCATGCTCATGTCGGCGCGCTTCGTG 436
Qy      GluGlyPheGlyValGlyIleLeuSerTyrThrValProValTyrIleAlaGluLeuSer 218
Db      437  GCGGCGATCGCGTGGGTGAGCTCTCATCATCGCGCGCGTGAACACGCGCGAGGTGTC 496
Qy      ProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIle 238
Db      497  CCGCGCTCGCGCGTGGGTCTCACATCTTCCGCGAGGTGTCTCACTCACTTCGCGCATC 556
Qy      PheLeuAlaTyrLeu-----LeuGlyMetPheIleProTTPArg 251
Db      557  CTCCTCGGATGATCTCAACTTCGCTTCGCCGCGCTCTCCCTCGCGCTCGCGCTGCGCG 616
Qy      LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPheIle 271
Db      617  ATTATGCTCGGCATAGCGCGGTGCGTCTCGCTCGCTTCATGCTGCTGCTGCTGCGCATG 676
Qy      ProGlnSerProArgTTPLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu 291
Db      677  CCGAGTCTCCCGGTGGCTCGTCATGAAGGCGGTCTCGCGACGCGCAAGGTGTGCTT 736
Qy      GlnValLeuArgGlyPheGluThrAspIleThrGluValAsnAspIleLysArgAla 311
Db      737  GCCAAGCGTCCGACCGCGGAGGCGCGGCGCATCGCGACATTAGACTGCC 796
Qy      312 Val-----Ala 313
Db      797  GCGGCGATCCCTCTGCGGCTTCGACGCGCGTGTGCTCCCGTCCCGCAAAACAAAGAGG 856
Qy      SerSerSerLysArgThr-----ThrIleSerPheGlnGlu 325
Db      857  AGCAGAGAGAGCGGTGTGAAGGACCTCATCTGTACCGACCATAGCATCGCCAC 916
Qy      LeuAsnGlnLysTyrArgThrProLeuLeuGlyIleGlyLeuValLeuGln 345
Db      917  ATC-----CTCATCGCGGAATCGGCATCCATCTTCTCCAG 952
Qy      346 AsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGly 365
Db      953  CAGTCTTCGGGCATPCGACGCGCGTGTCTAGCGCGGTAGTATTTCAAGAGCGCGCG 1012
Qy      366 ValThrAsnSerAspLeu-----AlaThrCysSerLeuGlyAlaIleGlnValLeu 382
Db      1013  ATCAGCGGCGACACGCGTCTCCGCGGACACCGTGGCGTGGCGGCCACCAATACGGTC 1072
Qy      383 AlaThrGlyValThrThrTrpLeuLeuAspAlaGlyArgArgIleLeuIle 402
Db      1073  TTCATCTGTGGCCACCTTCTCTCGACCGCATCGCGCGCGCGGTGTGTGTGAC 1132
Qy      403 SerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPheLeuLysAsp 422
Db      1133  AGACGCGGCGCATGCTGCTCTTAGTGGGCTCGCGAGGGGCTCACCGTCATCAGC 1192
Qy      423 AsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGlyIle 442
Db      1193  CGCCACCGGACGAGAGATCACTGGGCCATC-----GTCCTGTGCTATCTTCTGCATC 1246
Qy      443 ValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProThrLeuMetMetSer 462
Db      1247  ATGECCTACGTGGCTTCTTCTCATCGGCTCGGCCCATCATCAGTGGGTGACACTCG 1306
Qy      463 GluIleLeuProValSerIleLysSerLeuGlySerIleAlaThrLeuAlaAsnTrp 482
Db      1307  GAGATCTTCGCGTGCACGTCGCGGCTGGGTGTCTCCCTGCGCGGTGCGCGCAACCCG 1366
Qy      483 LeuThrSerPheAlaIleThrMetThr-----ThrAsnLeuMetLeuThrTrpSerValGly 501

```

```

Db      1367  CTGACCAAGCGCGTGTATCTCCATGACCTTCACTTCGCTGTCTCAAGGCATGACCATCGGC 1426
Qy      GlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValValLeuTrp 521
Db      1427  GCGCGCTTCTTCTTCTTCGCGCGCATGCGCTCATTCGATCGGTTCTTCTTCTTCTTCTT 1486
Qy      522 ValProGluThrLysGly 527
Db      1487  CTGCGGAGACCCGCGGC 1504

RESULT 4
US-09-291-922-21
; Sequence 21, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-291-922-21

Alignment Scores:
Pred. No.:      7,33e-51      Length:      2017
Score:          557.00      Matches:      148
Percent Similarity: 49.01%      Conservative: 99
Best Local Similarity: 29.37%      Mismatches: 211
Query Match:      20.04%      Indels:      46
DB:              3          Gaps:      9

US-10-051-909-36 (1-553) x US-09-291-922-21 (1-2017)
Qy      80 SerSerPheAlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSer--- 98
Db      75 TCTTCTCTGCACCACCGGAGATGCTTCGCCCGCTGCGCCATCTCGCTCATGACCTCATC 134
Qy      99 -----AlaPheLeuCysThrLeuIleValAlaLeuGlyProIle 111
Db      135 AAGAAGGCAACGTCGCGTTCGCTTCGCTGCGCCATCTCGCTCATGACCTCATC 194
Qy      112 GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu 131
Db      195 CTCCTCGCTACGATATCGGGGTGATGAGCGGGCGTGGCTGTGTACATCAAGAGGATTC 254
Qy      132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal 151
Db      255 AACATGATGACGGGAAGGTGAGGTTCCTCATGGGCATCTGAACCTCTACTCGCTCATC 314
Qy      152 GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIle 171
Db      315 GCCTCTTCGCGGCGGCGACGCTCGACTGGATCGCGCGGGGTACCATCGTGTTC 374
Qy      172 AlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPhe 191
Db      375 GCCCGCGTCATATCTTCGCGGGGGGTTCCTCATGGGGTTCGCGCACTACGCCATG 434
Qy      192 LeuTyrMetGlyArgLeuLeuGluPheGlyValGlyIleIleSerTyrThrValPro 211
Db      435 CTCATGTTCCGCCCTTCGTCGCGCGGATCGGCGGTACGCGCTCATGATCGCGCG 494
Qy      212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231

```

Db 495 GTGTACACCGCGAGGTGTGCGCGGCTGCGCGGTCTTCTGACGTGCTTCCCGGAG 554
Qy 232 LeuSerValThrPheGlyIlePheLeuAlaTyrLeu-----Leu 244
Db 555 GTGTTCATCACTTCGCGATCTCTGCGGTACGTCTCGAACTATGCTTCTCCCGCTTG 614
Qy 245 GlyMetPheIleProTArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeu 264
Db 615 CCGCTGAACCTCGGTGCGCATCATCTCGGCATCGCGCGCGTCCGTGCTC 674
Qy 265 IleProGlyLeuPhePheIleProGluSerProArgTrpLeuAlaIysMetAsnLeuThr 284
Db 675 GCGCTCATGTGCTGCGCATCGCGAGTCCCGCGTGTGCTGCTATGAAGGACGCCTC 734
Qy 285 GluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrGlu 304
Db 735 CCGGACCCCAAGGTGCTCTGGAGAACCTCCGACGCGGAGGCGCGGAGCGC 794
Qy 305 ValAsnAspIleIysArgAla-----ValAlaSer 314
Db 795 CTGGCCGACATCAAGCGCGCGCGCATCCCTGAGGAGCTCGACGCGGACGTGTGAC 854
Qy 315 SerSerIysArgThrThrIleSerPheGlnGluLeuAsnGlnIysLysTyrArgThrPro 334
Db 855 GTCCCCAAGAGGCGGAGCGGAACAGAGCGGTGTGAAGGAGCTCATCTGTCTCCCG 914
Qy 335 -----LeuLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeu 347
Db 915 ACCCGGCCATCGCGCATCTCTGTCGGGATCGGCATCCACTTCTTCCAGCATGCG 974
Qy 348 SerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheIysAlaIleGlyValThr 367
Db 975 TTGGGCATTCATCTCGTCTCTACAGCCCTCTGCTGTCTCAAGAGCCCGGATTACG 1034
Qy 368 AsnSerAspLeu-----AlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThr 384
Db 1035 AACGACAAACATCTTTGGGCACCATCTGCGCGTCTGCTGTCACCAAGAGCTTTTCATC 1094
Qy 385 GlyValThrThrTrpLeuLeuAspArgAlaGlyArgIleLeuLeuIleSerThr 404
Db 1095 TTGTGCGCACTTTCTTCATCGACGCGGTGCGCGCGCGCTGTGCTGGCGACGACG 1154
Qy 405 SerGlyMetThrLeuCysLeuLeuAlaValSerValIlePhePheLeuIysAspAsnIle 424
Db 1155 GGCGGATATCTCTCTCATCGGCTCGCGCGCGGCTCACCGTC-----GTC 1205
Qy 425 SerGlnAspSerAsnSerTyrTyrIleLeuThrMet---IleSerLeuValGlyIleVal 443
Db 1206 GGCCAGCACCCGCGACGCAAGATACCTTGGGCCATCGGCCCTAAGCATCGGCTCCACCC 1265
Qy 444 SerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetSerGlu 463
Db 1266 GCCTACATCGCTTCTTCTCATCGGCTTGGCCCCATCACGTGGGTGTACAGCTCGGAG 1325
Qy 464 IleLeuProValSerIleLysSerLeuGlySerIleAlaThrLeuAlaAsnTrpLeu 483
Db 1326 ATCTTCGCTCCAGGTGCGCGCTGGGTGCTGCTCGCTCGCGTCCGCCCAACCGGCTC 1385
Qy 484 ThrSerPheAlaIleThrMetThr---ThrAsnLeuMetLeuThrTrpSerValGlyGly 502
Db 1386 ACCAGCGCGCTCATCTCATGACCTTCTGCTGCTGTCGAAGGCCATCACCATCGCGGC 1445
Qy 503 ThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValValLeuTrpVal 522
Db 1446 AGCTTCTCTCTACTCTCGGCATCGCGCGCTCGCTCGCTGGGTGTCTTCTACACCTACCTC 1505
Qy 523 ProGluThrIysGly***AsnSerArg-----GlyAspThrIlePhe 536
Db 1506 CCGGAGACCGCGCGCGAGCGCTGGAGGAGATGAGCAAGCTGTTCGGCGACACGCGCGCC 1565
Qy 537 ValSerLeuSer 540
Db 1566 GCCTCGGAATCA 1577

RESULT 5
US-09-023-655-1104
; Sequence 1104, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1104:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3915 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g183684
; US-09-023-655-1104

Alignment Scores:
Pred. NO.: 9,468-50 Length: 3915
Score: 551.50 Matches: 145
Percent Similarity: 48.95% Conservative: 87
Best Local Similarity: 30.59% Mismatches: 180
Query Match: 19.85% Indels: 62
DB: 4 Gaps: 13

US-10-051-909-36 (1-553) x US-09-023-655-1104 (1-3915)
Qy 94 GluSerHisValSerAlaPheLeuCys-ThrLeuIleValAlaLeuGlyProIleGlnPh 113
Db 254 GAAGGTCAACCACTGGGTCATCAATGCTCGAGAGATCAATAAGGAATTTATCAATAA 313
Qy 113 eGlyPheThrSerGlyPhe----- 119
Db 314 TGGCTACAACTGGGTCATCAATGCTCGAGAGATCAATAAGGAATTTATCAATAA 373
Qy 120 -----SerSerProThrGlnAspAlaMetValArgAspLeu---As 132
Db 374 AACTTTGACGGACAAGGGAATGCCCCACCCTCTGAGGTGCTGCTCACTCTCTGCTC 433
Qy 132 nLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGl 152

```

Db      434  CTTGTCTGTGGCCATATTTTC-----GTCCGGGGTATGATCG 472
Qy      152  YAlaIleAlaSerGlyClnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAl 172
Db      473  CTCCTTTTCGTGGAGCTCTTCGTCAACCGCTTTGGCAGCGCAATTCATGCTGATGT 532
Qy      172  aAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPhe-----AlaLysAspAl 189
Db      533  CAACCTGTTGGCTGCTACTGCTGCTCTTATGGGACTGTGTAAAGTAGCTAAGTCGGT 592
Qy      189  aSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrTh 209
Db      593  TGAATGCTGATCTGCTGGTGGCTGTGTTATGTGCTCTTCCTGGGACTCTGCACAGTTT 652
Qy      209  rValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerVa 229
Db      653  TGTCCCATGATGATTCGGAGATCTCGCTACTGCTGCGGGTGCCTTTGGCACTCT 712
Qy      229  lAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMet---PheI 248
Db      713  CAACAGCTGGCATCTGTTGGAATCTGTGGTGGCCAGATCTTTGGTCTGGAATTCAT 772
Qy      248  e-----ProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMe 263
Db      773  CTTGGGTCTGAAGAGCTATGGCGCTGCTACTGCTGGTCTTACCATCTCTCTCTATCT 832
Qy      263  tLeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeuAlaLysMetAsnLe 283
Db      833  ACAAGTGCAGCCCTTCATTTGCTGCTGAAAGTCCAGATTTTGTCT---ATTAACAG 889
Qy      283  uThrGluaspCysGluThrSerLeuGlnValLeuArgGlyPhe-----GluThrAspI 301
Db      890  AAAAGAGAG---GAGAAATCTAAGCAGATCTCCAGCGTGTGGGCGACCCAGGATGT 946
Qy      301  eThrThrGluValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThrThrI 321
Db      947  ATCCAGACATCCAGGAGATGAAGATGAGTGCAAGTGTCAAGAAAAGCAAGT 1006
Qy      321  eSerPheClnGluLeu---AsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGl 340
Db      1007  CACGTGCTAGAGCTTTTAGAGTGTCCAGCTACCGACCCCATCATCTATTTCCATGT 1066
Qy      340  yLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerI 360
Db      1067  GCTCAGCTCTCTCAGCAGCTCTCTGGATCAATGCTGTCTTATTAATCTCAACAGGA 1126
Qy      360  ePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGl 380
Db      1127  CTTCAAGGATGCAGGTCTTCAAGAGCCCATCTATGCCACCATCGCGCGGTGTGTTAA 1186
Qy      380  nValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLe 400
Db      1187  TACTATCTTACTGTAGTTCTCTATTCTTGGTGGAAAGCGCAGAGAGAGACTGTGCA 1246
Qy      400  uIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLe 420
Db      1247  TATCATAGGCTTGGAGGATGGCTTTTGTTCACGCTCATGACTGTTCTTTGTTATT 1306
Qy      420  uLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuVa 440
Db      1307  AAAGGATAAC-----TATATGGGATGAGCTTTGCTGTATGG 1345
Qy      440  lGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMe 460
Db      1346  GGCTATCTGTGCTTTGTAGCTTTTGAATTTGAAATGGACAGGCCCATTCCTCGTTAT 1405
Qy      460  tMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAl 480
Db      1406  TGTGGCGGAATCTTTCAGCCAGGCGGCCCGCCAGCTGCGATGGCAGTGGCGGCTCTC 1465
Qy      480  aAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerVa 500

```

```

Db      1466  CAACTGGACCTCCAACTTCTCTAGTCGGA-----TTGCTCTTCCCTC 1507
Qy      500  lGlyGlyThrPheLeuSerTyrMetValValSerAlaPheThr----- 514
Db      1508  CGCTGCTCACTATTAGGAGCCTACGTTTATTATCTTCCCGGCTTCTCATTAACCTT 1567
Qy      515  lIleValPheValValLeuTrpValProGluThrLysGly 527
Db      1568  CTGCGCTTTTACCTCTCTCAAAAGTCCCTGAGACCCGTCG 1607

RESULT 6
US-09-291-922-23
; Sequence 23, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 23
; LENGTH: 1853
; TYPE: DNA
; ORGANISM: Glycine max
US-09-291-922-23,

Alignment Scores:
Pred. No.:      2,51e-49      Length:      1853
Score:          542.50      Matches:      137
Percent Similarity: 50.64%      Conservative: 102
Best Local Similarity: 29.03%      Mismatches: 183
Query Match:      19.52%      Indels:       51
DB:              3          Gaps:         9

US-10-051-909-36 (1-553) x US-09-291-922-23 (1-1853)
Qy      93  ArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGln 112
Db      239  CGAAAAGCAACAAGATATGCTTTGCTGTGCTATGCTGGCTCCATGACTTCCATCTT 298
Qy      113  PheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValAtgAspLeuAsn 132
Db      299  CTTGTTATGATATTTGGAGTGATGAGTGAGCAGCCATATACATAAAGAGGACTGAA 358
Qy      133  LeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGly 152
Db      359  GTCTCGGACGAGCAATCGAGATCTGCTCGGAATCATCAACCTATACCTCTCTGATAG 418
Qy      153  AlaIleAlaSerGlyClnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAla 172
Db      419  TCATGCTCTCGCGCAGAACCTCGAGTAGTAGTCCCGCTTACAGATTGTTTTCGCC 478
Qy      173  AlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeu 192
Db      479  GGCACCATCTCTTTGTTCGGAGCACTTCTCATGGGTTTCTCCCAATATTCTCTTCTC 538
Qy      193  TyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValProVal 212
Db      539  ATGTTTGGCCGTTTCGTCGCTGGCATTTGGCATCGGCTACGCCCTCATGATACCCCGTC 598
Qy      213  TyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeu 232
Db      599  TACACCGCGAGGTCCTCCCGGCTCTCTCTGTTGGCTTCTCACTTCTCTCCCTGAGTA 658
Qy      233  SerValThrPheGlyIlePheLeuAlaTyrLeu-----LeuGly 245

```



```
Db 659 TTTATTATGAGGGATATTAATGGATACATCAAACTATGCTATTTTCGAGCTGACA 718
Qy 246 MetPheIleProTprArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIle 265
Db 719 CTAAGGTGGATGGGATGATGCTTGGAGTGGTGCAATACCTTCGGTACTCTCAACA 778
Qy 266 ProGlyLeuPhePheIleProGlySerProArgTprLeuAlaLys----- 280
Db 779 GTAGGAGTGTGGCATGCGGAGTCCCCAAGTGGCTTGTGTATGAGGGGTGGTTGGGA 838
Qy 281 -----MetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArg 295
Db 839 GAGCGCAAGAAAGTGTAAACAACTCAGACGAGCAAGAGAGAGAGAGAGAGAGAGAG 892
Qy 296 GlyPheGluThrAspIleThrGluValAsnAspIleLysArgAlaVal-AlaSerSe 315
Db 893 -----AGCTAGCGGAATCAACAAGCCGACGAGGATCCCC 928
Qy 315 rSerLysArgThrIleSerPheGlnGluLeuAsnGln----- 328
Db 929 GAGAGTTGCAACGACGAGCTGTTAGGTAATAATAAACAAGCAACGAGTGTATGCG 988
Qy 329 -----LysLysTyrArgThrPro-----LeuLeuGlyIleG 340
Db 989 AAGAGCTCTCTCCT-TATCCAAAGCCGCAATTCGTACATCGTAATCGCTGCCCTGG 1047
Qy 340 YLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSer 360
Db 1048 TATTCACCTCTTCCAAACAGCGTGGGCGTAGAGCGCGTCTGTTTGTACAGCCCCAG 1107
Qy 360 ePheLysAlaAlaGlyValThrAsnSerAsp-----LeuAlaThrCysSerLeuG 377
Db 1108 CTTGCAAAAGCGTGGGATTAACAACGACGACGATAAGCTTCCTGCAACCGTGGCCCT 1167
Qy 377 YAlaIleGlnValLeuAlaThrGlyValThrTprLeuLeuAspArgAlaGlyArgAr 397
Db 1168 ATTCGTTAAGCCGTGTTATCTTGGCGGTACGTTTACGTTGACCGGTGGTCTGTCG 1227
Qy 397 gIleLeuLeuIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValVa 417
Db 1228 TCGGTTTATTGTCTAGTTCGCGGCGATGCTCTCGCTTCTCAGCGTTCGATCAG 1287
Qy 417 lPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMet 437
Db 1288 CTTCACTGTTATTGATCAT---TCGGAGAGGAATTAATGTCGGCGGTGGATCGAGCAT 1344
Qy 437 eSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIlePr 457
Db 1345 AGCCATCGTG-----TTGGCTTACGTGGCCACGTTCTCCATCGGTGGGTCCTCCAT 1398
Qy 457 cTprLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlySerIleAl 477
Db 1399 GTGGGTCTATAGTCTGAGATCTTCCGTTGAGCTCGCGGCGAGGTGCGCGCGGG 1458
Qy 477 aThrLeuAlaAsnTprLeuThrSerPheAlaIleThrMetThr---ThrAsnLeuMet 496
Db 1459 AGTTCGGGTGAATAGGACCACTAGCGGTTGTCATCATGACITTTCTGTCCTCACTAG 1518
Qy 496 uThrTprSerValGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleVa 516
Db 1519 AGCCATCACTATTGGTGAGCTTCTTCTTTTATTGTGGCATTGCTACTTGGGTGGAT 1578
Qy 516 lPheValValLeuTprValProGluThrLysGly 527
Db 1579 ATTCTTTTACACCGTCTGTGCTGAGACCGCGGGA 1612

RESULT 7
US-09-591-025-8
; Sequence 8, Application US/09591025
; Patent No. 6303373
; GENERAL INFORMATION:
; APPLICANT: Bogan, Jonathan S.
; APPLICANT: Lodish, Harvey F.
```

```
; TITLE OF INVENTION: Method of Measuring Plasma Membrane
; FILE OF INVENTION: Targeting of GLUT4
; FILE REFERENCE: 0399.1210-004
; CURRENT APPLICATION NUMBER: US/09/591,025
; CURRENT FILING DATE: 2000-06-03
; PRIOR APPLICATION NUMBER: 60/154,078
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/138,237
; PRIOR FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 2592
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified GLUT4 containing myc tag sequences
US-09-591-025-8

Alignment Scores:
Pred. No.: 1,79e-47 Length: 2592
Score: 528.00 Matches: 131
Percent Similarity: 52.58% Conservative: 83
Best Local Similarity: 32.19% Mismatches: 165
Query Match: 19.00% Indels: 28
DB: 3 Gaps: 7

US-10-051-909-36 (1-553) x US-09-591-025-8 (1-2592)
Qy 133 LeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGly 152
Db 562 CTCCTCGGCGCATCTTTTTC-----GTGGCGGCATGATTTCC 600
Qy 153 AlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAla 172
Db 601 TCCCTTCCTCATTTGGTATCATCTCTCAGTGGTGTGAAGAAAGGCGCATGCTGCTCAAC 660
Qy 173 AlaIleProAsnIleIleGlyTprLeuAlaIleSerPheAlaLysAspAlaSer----- 190
Db 661 AATGTCCTGGCGGTCTCGGGGCGAGCTCATGGCGCTGGCCCAACGCTGCTCCCTCAT 720
Qy 191 ---PheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThr 209
Db 721 GAAATGCTCATCTTGGACGATTCCTCATTTGGCGCTACTCAGGCTCATCATCGGGGTG 780
Qy 210 ValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerVal 229
Db 781 GTGCCCCATGATGCTGGGAGATTCCTCCACTCCTCCGCGGCGCCCTGGGAGCGCTC 840
Qy 230 AsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyr-----Leu 243
Db 841 AACCAACTGGCCATTGTTATCGGCATTCTGATCGCCAGGTGCTGGGCTTGGAGTCCCTC 900
Qy 244 LeuGlyMetPheIleProTprArgLeuLeuAlaValIleGlyAlaLeuProCysThrMet 263
Db 901 CTGGGCACTGCCAGCCTGTGGCCACTGCTCTGGGCGCTCAGAGTCTACCTGCCCTCTG 960
Qy 264 LeuIleProGlyLeuPhePheIleProGluSerProArgTprLeu---AlaLysMetAsn 282
Db 961 CAGCTGGTCTGCTGCCCTTCTGTCGCCGAGAGCCCGCTACCTCTACATCATCCAGAAT 1020
Qy 283 LeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGlu-----ThrAsp 300
Db 1021 CTCGAGGGGCTTGCACAGAAAGATCTGAAGCGCTGACAGGCTGGCGCGCATTTCTGGA 1080
Qy 301 IleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThrThr 320
Db 1081 GTGCTGGTGTAGCTGAAGATGAGAGCGGAGCTG-----GACGCTGAGCGGCATG 1134
Qy 321 IleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuGlyIleGly 340
Db 1135 TCCCTGCTCCAGCTCTCTGGGAGCGGTACCCAGCCGCGAGCCCTGATCATTCGCGTCTG 1194
```

Qy 341 LeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIle 360
 Db 1195 CTGCAGCTGAGCCAGCAGCTCTCTGGCATCAATGCTGTTTCTATTATTCAGCAGCATC 1254
 Qy 361 PheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGln 380
 Db 1255 TTCGACAGACAGGGGAGCCAGCCTGCTATGCCACCATAGAGAGCTGTGGTGGTCAAC 1314
 Qy 381 ValLeuAlaThrGlyValThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeu 400
 Db 1315 ACAGTCTTACCTTGGTCTCGGTGTTGTTGGAGCGGGCGGGCGGAGCCTCCAT 1374
 Qy 401 IleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeu 420
 Db 1375 CTCCTGGGCTGGCGGCGATGTTGGCTGTGCCATCTGATGACTGTGGCTCTGCTCTG 1434
 Qy 421 LysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuVal 440
 Db 1435 CTGGAGCGAGTTCCA-----GCCATGAGCTACGTTCCATTGTG 1473
 Qy 441 GlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMet 460
 Db 1474 GCCATCTTGGCTGCGTGGCATTTTGTAGATTGGCCCTGGCCCATTCCTTGGTTTCATC 1533
 Qy 461 MetSerGluIleLeuProValSerIleLysSerLeuGlySerIleAlaThrLeuAla 480
 Db 1534 GTGGCCGAGCTCTTCAGCCAGGAGACCCCGCGGAGCCATGCTGTGGTGTCTCTCC 1593
 Qy 481 AsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerVal 500
 Db 1594 AACTGGACGAGCAACTTCATCATTTGGCATGGGTTTCCAGTATGTCGGAGGCTATGGGG 1653
 Qy 501 GlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValValLeu 520
 Db 1654 CCCTACGCTCTCTCTTATTTCGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTTA 1713
 Qy 521 TrpValProGluThrLysGly 527
 Db 1714 AGAGTACTGAACTCGAGGC 1734

RESULT 8

US-09-291-922-19
 ; Sequence 19, Application US/09291922
 ; Patent No. 6383776
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Hitz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB-1163
 ; CURRENT APPLICATION NUMBER: US/09/291,922
 ; CURRENT FILING DATE: 1999-04-14
 ; EARLIER APPLICATION NUMBER: 60/083,044
 ; EARLIER FILING DATE: April 24, 1998
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 19
 ; LENGTH: 1914
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 US-09-291-922-19
 Alignment Scores:
 Pred. No.: 1,19e-47 Length: 1914
 Score: 527,50 Matches: 132
 Percent Similarity: 48.72% Conservatives: 97
 Best Local Similarity: 28.09% Mismatches: 186
 Query Match: 18.98% Indels: 55
 DB: 3 Gaps: 7
 US-10-051-909-36 (1-553) x US-09-291-922-19 (1-1914)

Qy 99 AlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGly 118
 Db 116 GCCTCCATATGTGCATCTCTGGCTCCATGCTCTGTCTCATCTTGTGCTATGACATTGGG 175
 Qy 119 PheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPhe 138
 Db 176 GTGATGATGGAGCGGCATGATCATCAAGAGGACCTGATATCAGGAGCGTCAAGCTG 235
 Qy 139 SerAlaPheGlySerLeuSerAsnValGlyMetValGlyAlaIleAlaSerGlyGln 158
 Db 236 GAGATCCTGATCGGATCTCTAGTCTCTACTCTGCTTTCGGATCTCTGCTGGCGCGG 295
 Qy 159 MetAlaGluTyrIleGlyArgLysSerLeuMetIleAlaAlaIleProAsnIleIle 178
 Db 296 ACGTCCGACAGATCGGCGCGCTTACCGCTGTGTTCGGCTGTCTCTCTCTCTCTCT 355
 Qy 179 GlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeu 198
 Db 356 GGCTGCTGTCTCATGGGTTTCGCCGTCACTACGCGCATGCTCATGCGCGCGCTTCGTG 415
 Qy 199 GluGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSer 218
 Db 416 GCCGAGTGGTGTGGGTACGCGGCGCATGATCGCGCCGCTGTACACGCGCGAGATCG 475
 Qy 219 ProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIle 238
 Db 476 CTGCGGCGTCCGTGGCTTCTTACACCTTCCCGGAGGTGTCTACATCAATCGGATC 535
 Qy 239 PheLeuAlaTyrLeu-----LeuGlyMetPheIleProTrpArg 251
 Db 536 CTGCTTGGCTACCTGTCCAACTTCGCGGCGCTCCCGCTCCACCTCGGCTGGCGC 595
 Qy 252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIle 271
 Db 596 GTCATGCTCGCATGGCGCGAGTTCCTGTCGGCTCTCTCGCGCTCTCTCTCTCTCTCT 655
 Qy 272 ProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu 291
 Db 656 CCCGAGTCCCTCGTGGTGGTCTTGAAGGGCGCGCTCGCGGACGCGGCTGTGCTA 715
 Qy 292 Gln----- 292
 Db 716 GAGAGACCTCTGCCACGCCAGAGAGCGCGCGAGCGGTGGCGCATCAAGCGCGCG 775
 Qy 293 -----ValLeuArgGlyPheGluThrAspIleThrThr----- 303
 Db 776 CGCGGATTCGAGAGGCGCTCGACGCGGAGCTAGTACCGTACCGCGCAGGAGCAAGGC 835
 Qy 304 -----GluValAsnAspIleLysArgAlaValAlaSerSerLysArgThrThrIle 321
 Db 836 GCGGTGAGTGTGAGGTGTGGAAGAGCTCATCTGTCTCCCGACCCCGCGGTGTG- 889
 Qy 322 SerPheGlnGluLeuAsnGlnLysTyrArgThrProLeuLeuLeuGlyIleGlyLeu 341
 Db 890 -----CGACGATCTCTCTCTCGCGCGGTGCTC 919
 Qy 342 LeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePhe 361
 Db 920 CACTTCTTCCAGCAGGCTTCTGGCAGCGACTCTCTCTCTCTCTCTCTCTCTCTCTCT 979
 Qy 362 LysAlaAlaGlyValThrAsnSerAspLeu-----AlaThrCysSerLeuGlyAla 378
 Db 980 AAGAGCGGGGATCACCGACGACCAAGCTCTCGGCGCTACCTCGCGGGTGGCGGTG 1039
 Qy 379 IleGlnValLeuAlaThrGlyValThrTrpLeuLeuAspArgAlaGlyArgArgIle 398
 Db 1040 ACCAAGAGCTTCTCATCTGTGGTGGCGACGCTCTCTCTCTCTCTCTCTCTCTCTCT 1099
 Qy 399 LeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValPhe 418
 Db 1100 CTGCTGTGATCAGCAGCGGCGGATGATTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1159
 Qy 419 PheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSer 438

```

Db 1160 ACGTCGCGGGGATACCCGACACCAAGTCGCGGCGGTC-----GCCCTGTGC 1213
Qy 439 LeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrp 458
Db 1214 ATCGCGTCAACCGTGTCTACATCGCGCTTCTTCCATCGCGCTCGGGCCCATACGGGC 1273
Qy 459 LeuMetSerGluIleLeuProValSerIleIysSerLeuGlyGlySerIleAlaThr 478
Db 1274 GTGTACACCTCGGAATATTCCCGCTGACAGTGGCGCGCTGGCTTCGGGGTGGTGG 1333
Qy 479 LeuAlaSerTrpLeuThrSerPheAlaIleThrMetThr---ThrAsnLeuMetLeuThr 497
Db 1334 GCGAGCAACCGCGTACCAGCGCGTCACTCCATGACCTTCCTGCTCCCTCCCAAGGCC 1393
Qy 498 TrpSerValGlyIleThrPheLeuSerTrpMetValValSerAlaPheThrIleValPhe 517
Db 1394 ATCACCATCGCGCGACCTTCTCTCTACTCTCGGCATCGCGGTCGCGGTTC 1453
Qy 518 ValValLeuTrpValProGluThrLysGly 527
Db 1454 TTCTTACGTGCTCCCGGAGACACGCGGC 1483

```

RESULT 9

```

US-09-894-927B-8
; Sequence 8, Application US/09894927B
; Patent No. 6632924
; GENERAL INFORMATION:
; APPLICANT: Bogan, Jonathan S.
; APPLICANT: Lodish, Harvey F.
; TITLE OF INVENTION: Method of Measuring Plasma Membrane
; FILE REFERENCE: 0399.1210-005
; CURRENT APPLICATION NUMBER: US/09/894,927B
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 09/591,025
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: US 60/154,078
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: US 60/138,237
; PRIOR FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 2592
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified GLUT4 containing myc tag sequences
; NAME/KEY: CDS
; LOCATION: (1)...(2592)
US-09-894-927B-8

```

```

Alignment Scores:
Pred. No.: 4.94e-47 Length: 2592
Score: 524.00 Matches: 131
Percent Similarity: 52.33% Conservative: 82
Best Local Similarity: 32.19% Mismatches: 166
Query Match: 18.86% Indels: 28
DB: 4 Gaps: 7

```

US-10-051-909-36 (1-553) x US-09-894-927B-8 (1-2592)

```

Qy 133 LeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyIleMetValGly 152
Db 562 CTCCTCCGTGGCCATCTTTCC-----GTGGCGGCATGATTCC 600
Qy 153 AlaIleAlaSerGlyGlnMetAlaGluTrpIleGlyArgLysGlySerLeuMetIleAla 172
Db 601 TCCTTCTCATGTGATCACTCTCAGTGGCTTGGAGAGAAAGGCCCATGCTGGTCAAC 660
Qy 173 AlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaIysAspAlaSer----- 190

```

```

Db 661 AATGCTCTGGCGGTCTGCTGGGGGAGCCTCATGGSCCTGGCAACGCTGCTCCTCTAT 720
Qy 191 ---PheLeuTyMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTrpThr 209
Db 721 GAAATGCTCATCTCTTGGACGATTCTTCATTTGGCGCTACTCAGGCTCAGCATCAGGGTG 780
Qy 210 ValProValTyIleAlaGluIleSerProGlnAsnMetAspGlyAlaLeuGlySerVal 229
Db 781 GTGCCCATGTAGTGGGGGAGATTGCTCCACTCACCCTCGGGGGGCGCTGGGAGGCTC 840
Qy 230 AsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyr-----Leu 243
Db 841 AACCAACTGGCCATTGTTATCGGCATTCTGATCGCCAGCTGCTGGGCTTGGAGTCCCTC 900
Qy 244 LeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMet 263
Db 901 CTGGGCACTGCCAGCGCTGGCCACTGCTCTGGGCCCTCAGAGTCTACCTGCTCCCTCTG 960
Qy 264 LeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeu---AlaLysMetAsn 282
Db 961 CAGCTGGTCTGCTGCTCCCTTCTGTCGCCGAGAGCCCGCTACCTCTACATCATCCAGAT 1020
Qy 283 LeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGlu-----ThrAsp 300
Db 1021 CTCGAGGGCGCTGCCAGAAAGAGTCTGAAGCGCCTGACAGGCTGGCCGATGTTCTGGA 1080
Qy 301 IleThrThrGluValAsnAspIleIysArgAlaValAlaSerSerSerIysArgThrThr 320
Db 1081 GTGCTGGTGTAGCTGAAGATGAGAAGCGAAGCTG-----GAGCGTGAGCGGCACCTG 1134
Qy 321 IleSerPheGlnGluLeuAsnGlnIysIysTyArgThrProLeuLeuLeuGlyIleGly 340
Db 1135 TCCTTGCTCCAGCTCTCGGCAGCGCTACCCAGCGGCGCCCTGATATCGCGTGTG 1194
Qy 341 LeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyAlaSerSerIle 360
Db 1195 CTGACGTGAGCCAGCAGCTCTCTGGCATCATGCTGTTTCTATTATTCGACCGATC 1254
Qy 361 PheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGln 380
Db 1255 TTCGAGACAGCAGCGGTAGCGCCTGCTATGCCACCATAGGAGCTGTGTGTCTCAAC 1314
Qy 381 ValLeuAlaThrGlyValThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeu 400
Db 1315 ACAGTCTTTCACCTTGTCTCGGTGTGTGGTGGAGCGGCGCGCGCGCGCTCCAT 1374
Qy 401 IleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeu 420
Db 1375 CTCTCGGGCTGCGGGCATGTGTGCTGTGCATCTGATGACGTGGCTCTGTCTCTG 1434
Qy 421 LysAspAsnIleSerGlnAspSerAsnSerTyTyIleLeuThrMetIleSerLeuVal 440
Db 1435 CTGGAGCGAGTTCCA-----GCCATGAGCTAGCTCTCCATTGTG 1473
Qy 441 GlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMet 460
Db 1474 GCCATCTTGGCTTGTGGCATTTTGTAGATTGGCCCTGGCCCATCTCTTGGTTCATC 1533
Qy 461 MetSerGluIleLeuProValSerIleLysSerLeuGlyIysSerIleAlaThrLeuAla 480
Db 1534 GTGGCGAGCTCTTCAGCGAGGACCCCGCGGAGCCATGCTGTGGCTGTCTCTCC 1593
Qy 481 AsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerVal 500
Db 1594 AACTGACGAGCAACTTCATTCATTGGCATGGGTTTCCAGATGTTTGGGAGGCTATGGGG 1653
Qy 501 GlyIleThrPheLeuSerTyMetValValSerAlaPheThrIleValPheValValLeu 520
Db 1654 CCTAGTCTCTCTCTATTTCGGCTCTCTCTGCTGCTCTCTCTCTCTCTCTCTCTCT 1713
Qy 521 TrpValProGluThrLysGly 527

```

Db 1714 AGAGTACCTGAAGCTCGAGC 1734

RESULT 10

US-09-643-597-135

Sequence 135, Application US/09643597

Patent No. 6426072

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong

APPLICANT: Fan, Liqun

APPLICANT: Kalos, Michael D.

APPLICANT: Bangur, Chaitanya S.

APPLICANT: Hosken, Nancy

APPLICANT: Fanger, Gary R.

APPLICANT: Li, Samuel X.

APPLICANT: Wang, Aijun

APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Henderson, Robert A.

APPLICANT: McNeill, Patricia D.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.455C11

CURRENT APPLICATION NUMBER: US/09/643,597

CURRENT FILING DATE: 2000-08-21

NUMBER OF SEQ ID NOS: 369

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 135

LENGTH: 2856

TYPE: DNA

ORGANISM: Homo sapien

US-09-643-597-135

Alignment Scores:

Pred. No.: 6.64e-47 Length: 2856

Score: 523.50 Matches: 151

Percent Similarity: 45.91% Conservative: 96

Best Local Similarity: 28.07% Mismatches: 229

Query Match: 18.84% Indels: 62

DB: 4 Gaps: 12

US-10-051-909-36 (1-553) x US-09-643-597-135 (1-2856)

Qy 27 ProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlySer-AsnArgG1 46

Db 14 CCGAGTGCAGCAGCCGAGGAGCAGAGACCAACGACGGGGTGCAGAGTCAGAGTCGAG 73

Qy 46 YGlyAla-----GlyAlaGlyGluSerGlySerAspHisAspGlyValLeuAr 63

Db 74 TGGAGTCCCGGACCGAGCAGAGCCTGACGGGAGAGCGCGCTCGCAGCCCGCTCG 133

Qy 63 gArgProLeuLeuAsnThrGlySerTrpTyrArgMetSerSerArgGlnSerSerPheAl 83

Db 134 CCACCCCGTACCCGG-----CGCAGCAGAGCCACCCAGCGAGCCGCTGC 178

Qy 83 aProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCyth 103

Db 179 CATGGAGCCAGCAGCAGAGCTGACGGGTCC-----CTCATGTGCTGT 226

Qy 103 rLeuileValAlaLeuGlyProLeinPheGlyPheThrSerGlyPheSerSerProTh 123

Db 227 GGGAGGAGCAGTGTCTGGCTCCCTGCAGTTTGGCTACACACTGGAGTCATCAATGCCCG 286

Qy 123 rGlnAspAlaMetValArgAspLeuAn----- 132

Db 287 CCAGAGGTATCGAGGAGTTCTACACACAGACATGGTCCACCGCTATGGGAGAGCAT 346

Qy 133 -----LeuSerIleSerGluPheSerAlaPheG1 142

Db 347 CCTGCCACACAGCTCACCAGCTCTGGTCCCTCAGTGGCCATCTTTCT----- 398

Qy 142 ySerLeuSerAsnValGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTy 162

Db 399 -----GTTGGGGCATGATGGCTCTCTCTCTGGGGCTTTTCGTTAACCG 445

Qy 162 rIleGlyArgGlySerGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAl 182

Db 446 CTTTGGCCGGCGGAATTCATGCTGATGATGAACCTGCTGCCTTCGTGTCGGCGTGTCT 505

Qy 182 aIleSerPheAla-----LysAspAlaSerPheLeuTyrMetGlyArgLeuLeuG1 199

Db 506 CATGGCTTCTCGAAACTGGGCAAGTCTTTGAGATGCTGATCTCGTGGCGCTTCATCAT 565

Qy 199 uGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPr 219

Db 566 CGGTGTGTACTCGCGCCCTGACACAGAGCTTCGTCCTCATGTATGTGGTGAAGTGCACC 625

Qy 219 oGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePh 239

Db 626 CACAGCCCTTTCGTGGGGCCCTGGGCAACCTGCACAGCTGGGCATCTCGTCGGCATCT 685

Qy 239 eLeuAlaTyrLeuLeuGlyMetPheIle-----ProTrpArgLeuLe 253

Db 686 CATGCCCGAGGTGTTCGGCCCTGGACTCCATCATGGGCAACAGGACCTGTGGCCCTGTCT 745

Qy 253 uAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProG1 273

Db 746 CTTGAGCATCATCTTCATCCCGCCCTGCTGCAGTGCATCGTGTGCCCTTCCTCCCGA 805

Qy 273 uSerProArgTrpLeuAlaLysMetAsnLeuThrGluAsp-----CysGluThrSerLe 291

Db 806 GAGTCCCGCTTCTCTGCTC---ATCAACCGCAACGAGGAGAACCGGCGCCAGAGTGTCT 862

Qy 291 uGlnValLeuArgGlyPheGluThrAspIleThrGluValAsnAspIleLysArgAl 311

Db 863 AAGAAGCTGCGCGG---ACAGCTGAGTGCACCATGACCTGCAGAGAGTGAAGGAGA 919

Qy 311 aValAlaSerSerLysArgThrThrIleSerPheGlnGluLeu---AsnGlnLysLy 330

Db 920 GAGTCGGCAGATGATGCGGGAGAGAAGGTCAACCATCTCGAGAGTGTTCGCTCCCGCG 979

Qy 330 sTyrArgThrProLeuLeuGlyIleGlyLeuValLeuGlnAsnLeuSerGlyI1 350

Db 980 CTACCGCAGCCCATCTCTCATCGTGTGGTGTGCAGCTGCCAGCAGCTGTCTGCAT 1039

Qy 350 eAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAs 370

Db 1040 CAACGCTGTCTCTATCTTACTCCACGAGCATCTTCGAGAGCGGGGTGCAGCAGCCTGT 1099

Qy 370 pLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrpLe 390

Db 1100 GTATGCCACCATGGCTCCGTATCGTCAACACGCGCTTCACTGCTGTGCTGTGTGT 1159

Qy 390 uLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCy 410

Db 1160 GTTGAGCGAGCAGCGCGGACCTTCGACCTCATAGGCTCGCTGGCATGGCGGTG 1219

Qy 410 sLeuLeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSe 430

Db 1220 TCCCATATCATCACTACCATCGGCTAGCAGCTGCTGGAGCAGCTACCTCTG----- 1268

Qy 430 rTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSe 450

Db 1269 -----ATGTCCTATCTGAGCATCGTGGCATCTTTGGCTTTTGGCTTCTTTGA 1318

Qy 450 rPheGlyMetGlyAlaIleProTrpLeuMetSerGluIleLeuProValSerIleLy 470

Db 1319 AGTGGTCTCTGGCCCATCCCATCGTTTCATCGTGGCTGAACCTCTACGCGAGGTCACG 1378

Qy 470 sSerLeuGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMe 490

Db 1379 TCCAGCTCCCATCGGTGAGGCTTCCAACTGGACCTCAAAATTCATTGTGGCAT 1438

Qy 490 tThrThrAsnLeuMetLeuThrTrpSerValGlyThrPheLeuSerTyrMetValVa 510

Db 1439 GTGCTTCCAGTATGTGGAGCACTGTGTGGTCCCTACGCTTTCATCATCTTCACTGTCT 1498

Qy 510 lSerAlaPheThrIleValPheValLeuTrpValProGluThrLysGly 527

```

Db      1499  CCTGGTCTCTGTTCTTCTTCATCTTCACCTACTCTCAAAAGTTCCTGAGACTAAAGGC 1550
          :          :          :          :          :          :          :          :          :          :
RESULT 11
US-09-480-884A-135
; Sequence 135, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Ligu
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Pangeri, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-480-884A-135

Alignment Scores:
Pred. No.: 6,648-47      Length: 2856
Score: 523.50      Matches: 151
Percent Similarity: 45.91%      Conservative: 229
Best Local Similarity: 28.07%      Mismatches: 96
Query Match: 18.84%      Indels: 62
DB: 4      Gaps: 12

US-10-051-909-36 (1-553) x US-09-480-884A-135 (1-2856)

Qy      27  ProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlyGlySer-AsnArgEl 46
Db      14  CCGAGTGCAGCCGCCAGGAGGAGCAGGAGACCAACGACGGGGTCCGAGTCAGAGTCGCCAG 73
Qy      46  yGlyAla-----GlyAlaGlyGluSerGlySerAspHisAspGlyValleuAr 63
Db      74  TGGAGTCCCGGACCGGAGCAGAGCTGAGCGGAGGAGCGCCGCTCGCAGCCCGCTGC 103
Qy      63  gArgProLeuLeuAsnThrGlySerTrpTyrArgMetSerSerArgGlnSerSerPheAl 83
Db      134  CCACCCGCTACCGG-----CGCAGCCAGAGCCACCAGCGCAGCGCTGC 178
Qy      83  aProGlyThrSerSerMetAlaValleuArgGluSerHisValSerAlaPheLeuCystH 103
Db      179  CATGGAGCCCGACGACGACCAAGACTCAGCGGTGC-----CTCANGCTGCTGT 226
Qy      103  rLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProth 123
Db      227  GGAGAGCAGCAGTCTTGGCTCCCTCGAGTTGGCTTACAACTGAGTGCATCATGCCCC 286
Qy      123  rGlnAspAlaMetValArgAspLeuAsn-----
Db      287  CCAGAAAGGTGATCGAGGAGTTTCTACAAACGACACATGGGTCCACCGCTATGGGAGAGCAT 346
Qy      133  -----
Db      347  CTTGCCACCAACGCTCACCACGCTTGGTCCCTCTCAGTGGCCATCTTTCT----- 398
Qy      142  ySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTy 462
Db      399  -----GTTGGGGGCGATGATTGGCTCTCTTCTCTGTGGGCCCTTTTTCGTTAACCG 445
Qy      162  rIleGlyArgLysGlySerLeuMetIleAlaIleProAsnIleIleGlyTrpLeuAl 192
Db      446  CTTTGGCCGCGCGAATCAATCATGATGAACTGCTGGCCCTTCGTGTCGCGCGGTCT 505
Qy      182  aIleSerPheAla-----LysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGl 199

```

```
; Sequence 135, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.45508
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-542-615A-135

Alignment Scores:
Pred. No.: 6,64e-47 Length: 2856
Score: 523.50 Matches: 151
Percent Similarity: 45.91% Conservative: 96
Best Local Similarity: 28.07% Mismatches: 229
Query Match: 18.84% Indels: 62
DB: 4 Gaps: 12

US-10-051-909-36 (1-553) x US-09-542-615A-135 (1-2856)
Qy 27 ProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlySer-AsnArgG1 46
Db 14 CCGAGTGCAGCAGCCAGGAGCAGGAGACCAACAGCGGGGCGGAGTCAGATCGCAG 73
Qy 46 yGlyAla-----GlyAlaGlyGluGluSerGlySerAspHisAspGlyValLeuAr 63
Db 74 TGGGAGTCCCGGACCGGAGCAGGAGCCTGAGCGGGAGAGCGCGCTCGCAGCGCGTGC 133
Qy 63 gArgProLeuLeuAsnThrGlySerTrpTyArgMetSerSerArgGlnSerSerPheAl 83
Db 134 CCACCCCGCTACCCGG-----CCGAGCCAGAGCACCAGCGGAGCGCTGC 178
Qy 83 aProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCysTh 103
Db 179 CATGGAGCCCGCAGCAGCAAGAAGCTGACGGGTGC-----CTCATGTGCTGT 226
Qy 103 rLeuileValAlaLeuGlyProileGlnPheGlyPheThrSerGlyPheSerSerProth 123
Db 227 GGGAGGAGCAGTGCTTGGCTCCCTGAGTTTGGCTACACACTGGAGTCATCATGCCCC 286
Qy 123 rGlnAspAlaMetValArgAspLeuAsn----- 132
Db 287 CCAGAAGGTATCGAGGAGTTCTACACACAGACATGGTCCACCGCTATGGGAGAGCAT 346
Qy 133 -----LeuSerIleSerGluPheSerAlaPheG1 142
Db 347 CTGCCCCACACGCTCACCAGCTCTGGTCCCTCTCAGTGGCCATCTTTCT----- 398
Qy 142 ySerLeuSerAsnValGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTy 162
Db 399 -----GTTGGGGGCATGATGGCTCTCTCTGTGGGCTTTTCGTTAACGC 445
Qy 162 rIleGlyArgGlySerLeuMetIleAlaIleProAsnIleIleGlyTrpLeuAl 182
Db 446 CTTTGGCCGCGGAATTCATATGATGATGAACCTCGTGGCCTTCGTCGCCCGTGT 505
Qy 182 aIleSerPheAla-----LysAspAlaSerPheLeuTyMetGlyArgLeuLeuG1 199
Db 506 CATGGGGTTCTCGAAACTGGGCAAGTCCCTTGAGATGCTGATCTCGGCGGCTTCATCAT 565
Qy 199 uGlyPheGlyValGlyIleSerTyThrValProValTyRileAlaGluIleSerPr 219
```

```
Db 566 CGGTGTGTACTCGGCGCTGCACAGGCTTCGTGCCCATGTATGTGGTGAAGTGTACC 625
Qy 219 oGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePh 239
Db 626 CACAGCCTTTGCTGGGCGCTGGGCAACCTGCACAGCTGGCATCGCTCGCGCATCTCT 685
Qy 239 eLeuAlaTyRleuLeuGlyMetPheIle-----ProTrpArgLeuLe 253
Db 686 CATGCCCCAGGTGTTCGCCCTGGACTCATCATCGGCAACAAGGACCTGTGGCCCTGTCT 745
Qy 253 uAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProG1 273
Db 746 GCTGAGCATCATCTTCATCCCGGCGCTGCTGCATGTCATCGTCTGCTCTTCGCCCGCA 805
Qy 273 uSerProArgTrpLeuAlaLysMetAsnLeuThrGluAsp-----CysGluThrSerLe 291
Db 806 GAGTCCCGCTTCTGTCTC-----ATCAACCGCAACAGGAGAACCGGCGCAAGAGTGTCT 862
Qy 291 uGlnValLeuArgGlyPheGluThrAspIleThrGluValAsnAspIleLysArgAl 311
Db 863 AAAGAAGCTGCGCGG---ACAGCTGAGCTGACCATCAGCTGCAGGAGATGAAGGAAGA 919
Qy 311 aValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeu---AsnGlnLysLy 330
Db 920 GAGTCGGCAGATGATCGGGAGAAAGAGTCCACCATCTCGAGAGCTGTCCGCTCCCCCGC 979
Qy 330 sTyArgThrProLeuLeuLeuGlyIleGlyLeuValLeuGlnAsnLeuSerGlyI1 350
Db 980 CTACCGCCAGCCATCTCTCATCGCTGTGTGCTGCTGACGCTGCCAGCAGCTGTCTGCAT 1039
Qy 350 aAsnGlyValLeuPheTyAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAs 370
Db 1040 CAACGCTGTCTTCTATTACTCCAGCAGCATCTTCAGAGCGGGGTGCAGCAGCCTGT 1099
Qy 370 pLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrpLe 390
Db 1100 GTATGCCACCATTCGCTCCGTATCTGCAACACGCGCTTCTACTGCTGTGCTGTTTGT 1159
Qy 390 uLeuAspArgAlaGlyArgGlyLeuLeuLeuIleSerThrSerGlyMetThrLeuCy 410
Db 1160 GGTGAGCGAGCAGCGCGCGGAGCCCTGCACCTCATAGGCTCTCGTGGCATGGCGGTG 1219
Qy 410 sLeuLeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSe 430
Db 1220 TGCCATATCTATGACCATCGCGCTAGCAGCTGCGAGCAGCTACCCCTGG----- 1268
Qy 430 rTyTyRileLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSe 450
Db 1269 -----ATGTCCTATCTGAGCATCGTGGCCATCTTTGGCTTTTGGCCTCTTTGA 1318
Qy 450 rPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLy 470
Db 1319 AGTGGGTCTCGGCCCATCCCATGTTTCATCGTGGTGAATCTCTACGCCAGGCTCCACG 1378
Qy 470 sSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMe 490
Db 1379 TCCAGCTGCCATTCGCGTTCAGCGCTTCTCCAACTGAGACCTCAAATTCATTGTGGCAT 1438
Qy 490 tThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyMetValva 510
Db 1439 GTGCTTCCAGTATGTGGAGCAACTGTGTGTGCTCCCTACGCTCTTCATCATCTTCACTGTCT 1498
Qy 510 lSerAlaPheThrIleValPheValValLeuTrpValProGluThrLysGly 527
Db 1499 CCGTGTCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1550

RESULT 13
US-09-606-421B-135
; Sequence 135, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
```


FILE REFERENCE: 210121.455C2
CURRENT APPLICATION NUMBER: US/09/221.107
CURRENT FILING DATE: 1998-12-22
NUMBER OF SEQ ID NOS: 161
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 135
LENGTH: 2856
TYPE: DNA
ORGANISM: Human
US-09-221-107-135

Alignment Scores:
Pred. No.: 6.64e-47 Length: 2856
Score: 523.50 Matches: 151
Percent Similarity: 45.91% Conservative: 96
Best Local Similarity: 28.07% Mismatches: 229
Query Match: 18.84% Indels: 62
Gaps: 12

US-10-051-909-36 (1-553) x US-09-221-107-135 (1-2856)

Qy	27	ProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlyGlySer-AsnArgG1	46
Db	14	CCGAGTGCAGCAGCCAGGAGCAGGAGACCAACAGCGGGGTCGGAGTCAGAGTCGAG	73
Qy	46	yGlyAla-----GlyAlaGlyGluGluSerGlySerAspHisAspGlyValLeuAr	63
Db	74	TGGAGTCCCGGACCGGAGCAGGAGCCTGAGCGGGAGAGCGCGCTCGCAGCCCGTCG	133
Qy	63	gAtgProLeuLeuAsnThrGlySerTrpTyArgMetSerSerArgGlnSerSerPheAl	83
Db	134	CCACCCCGTACCCGG-----CGAGCCAGGACACAGCCGCGAGCCTGC	178
Qy	83	aProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCysTh	103
Db	179	CATGAGCCCGCAGCAGCAAGAGCTGACGGGTGC-----CTCATGTGGCTGT	226
Qy	103	rLeuLeValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProth	123
Db	227	GGGAGGAGCAGTGTTCCTGCTCCCTGAGTTGGTCTACACACTGGAGTCATCAATCCCC	286
Qy	123	rGlnAspAlaMetValArgAspLeuAn-----	132
Db	287	CCAGAGGTATCGAGGAGTGTACACACAGACATGGTCCACCGCTATGGGGAGAGCAT	346
Qy	133	-----LeuSerIleSerGluPheSerAlaPheG1	142
Db	347	CCTGCCACACAGCTACACAGCTCTGTGCTCCCTCTCAGTGGCCATCTTTTCT	398
Qy	142	ySerLeuSerAsnValGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTy	162
Db	399	-----GTTGGGGGCGATGATGGCTCCTCTCTGTGGGCTTTTCGTTAAACCG	445
Qy	162	rIleGlyArgGlySerLeuMetIleAlaIleProAsnIleIleGlyTrpLeuAl	182
Db	446	CTTTGGCGCGGAATCAATGATGATGATGAACCTGCTGGCCCTCTGTGCTGGCCGTGT	505
Qy	182	aIleSerPheAla-----LysAspAlaSerPheLeuTyMetGlyValGluLeuG1	199
Db	506	CATGGGCTTCTCGAACTGGGAGTCTTTGAGATGCTGATCTGGGCGCTTCAATCAT	565
Qy	199	uGlyPheGlyValGlyIleIleSerTyThrValProValTyIleAlaGluIleSerPr	219
Db	566	CGGTGTGTACTGGCGCTGACCAAGCTTCGTGGCCCATGATGGGTGAAGTGCACC	625
Qy	219	oGlnAsnMetArgAlaLeuGlySerValAnGlnLeuSerValThrPheGlyIlePh	239
Db	626	CACAGCCTTTCGTGGGCGCTGGGACCTGTCACAGCTGGCATCGTGTGGCATCTCT	685
Qy	239	eLeuAlaTyLeuLeuGlyMetPheIle-----ProTrpArgLeuLe	253
Db	686	CATGCCCGGAGTGTTCGGCTGGACTCCATCATGGGCAACAGGACCTGTGGCCCTGT	745

Qy	253	uAlaValIleGlyAlaLeuProCysThrMetLeuLeuProGlyLeuPhePheIleProG1	273
Db	746	GCTGAGCATCATCTTCCCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	805
Qy	273	userProArgTrpLeuAlaLysMetAsnLeuThrGluAsp-----CysGluThrSerLe	291
Db	806	GAGTCCCGCTTCTGCTC-----ATCAACCGCAAGAGGAAACCGGCCCAAGAGTGTGT	862
Qy	291	uGlnValLeuArgGlyPheGluThrAspIleThrGluValAsnAspIleLysArgAl	311
Db	863	AAAGAGCTGCGGG-----ACAGCTGAGCTGACCTGACCTGAGGAGATGAAGGAAGA	919
Qy	311	aValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeu-----AsnGlnLysLy	330
Db	920	GAGTCGGCAGATGATGCGGAGAGAAAGGTACCATCTGAGAGCTGTTCGCTCCCCCGC	979
Qy	330	sTyArgThrProLeuLeuLeuGlyIleGlyLeuValLeuGlnAsnLeuSerGlyI1	350
Db	980	CTACCCGAGCCCATCTCTATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1039
Qy	350	eAsnGlyValLeuPheTyAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAs	370
Db	1040	CAACGCTCTCTTATTACTCCAGGAGCATCTTCGAGAAAGCGGGGTGCAGAGACCTGT	1099
Qy	370	pLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrpLe	390
Db	1100	GTATGCCACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1159
Qy	390	uLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCy	410
Db	1160	GCTGAGGAGCAGCGCGCGGAGCCCTGACCTCATAGGCTGCTGCTGCTGCTGCTGCTG	1219
Qy	410	sLeuLeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSe	430
Db	1220	TGCATATCATGACCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1268
Qy	430	rTyThrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSe	450
Db	1269	-----ATGTCCTATCTGAGCATCGGCCATCTTTGGCTTTGGGCTCTTTGA	1318
Qy	450	rPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluLeuLeuProValSerIleLy	470
Db	1319	AGTGGTCTCGCGCCCATCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1378
Qy	470	sSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMe	490
Db	1379	TCCAGTCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1438
Qy	490	tThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyMetValVa	510
Db	1439	GTGCTTCCAGTATGAGGAGCACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1498
Qy	510	lSerAlaPheThrIleValPheValLeuTrpValProGluThrLysGly	527
Db	1499	CCTGCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1550

RESULT 15
US-09-466-396A-135
Sequence 135, Application US/09466396A
Patent No. 6696247
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.455C4
CURRENT APPLICATION NUMBER: US/09/466.396A
CURRENT FILING DATE: 1999-12-17
NUMBER OF SEQ ID NOS: 224
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 135
LENGTH: 2856
TYPE: DNA
ORGANISM: Homo sapien

115 Page Blank (uspio)


```

; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 61061
; LENGTH: 2003
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_62530C.1
; US-10-437-963-61061

Alignment Scores:
Pred. No.:      2,42e-203      Length:      2003
Score:          2133.00        Matches:      434
Percent Similarity: 79.77%    Conservative: 47
Best Local Similarity: 79.20% Mismatches:    51
Query Match:      76.75%     Indels:      17
DB:              17          Gaps:         7

US-10-051-909-36 (1-553) x US-10-437-963-61061 (1-2003)
QY 1 ProSerSerSerSerSerPheArgProAlaGlyLysLysLysLysLysLysLysGlnGly 20
DB 55 CCTTCAATTCATCT-----CGTCTCTCTCTCGTAACTCG-----Val 38
QY 21 LeuArgArgGlnAlaValProGlyArgProAlaSerGlnLeuArgThrArg-----Val 38
DB 91 ---TAAAGAGATCTGCTCTCGCTCGGGAATCAACCATTTGCCGGCGAAGTAGAGATCG 147
QY 39 MetGlyGlySerSerArg---GlyGlyAlaGlyAlaGlyGlnGluSerGlySerArg 57
DB 148 GAGGGCGGGCCCATGAACAGGGCGGGCGGCGTCTCGGCGCATGAGAGCGGAACGAC 207
QY 58 HisAsp-----GlyValLeuArgArgProLeuLeu---AsnThrGlySerTyrTyrArg 74
DB 208 TAGAGAGCGGGCGGCATCGCGAGGCCCTCTCTGATGCACACGGGGAGCTGGTACAGG 267
QY 75 MetSerSerArgGlnSerSerPhe---AlaProGlyThrSerSerMetAlaValLeuArg 93
DB 268 ATGGGGTCGGCGAGGGGAGCCTCACCGGCGGGGACCTCGTCCATGGCCATCTTACGC 327
QY 94 GluSerHisValSerAlaPheLeuLeuCysThrLeuValAlaLeuGlyProGlnPhe 113
DB 328 GAGTCCACATCTCCCGCTTCCTCTGCACGCTCATGCTCGCGTCCGCCCCATCCAGTTC 387
QY 114 GlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeu 133
DB 388 GGATTCACGGGGGCTCTCTCTCCCGCAGCAGGAGCCATCATCCGAGACCTCGACCTC 447
QY 134 SerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAla 153
DB 448 ACCCTCCAGAGTCTCGGTGTTCCGATCGCTCTCCAAACGCTCGCGCCATGGTGGGGCG 507
QY 154 IleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAla 173
DB 508 ATTCAGAGTGTGATGCGGAGTACATTTGGCGGCAAGGGTCATTGATGATGCTGCA 567
QY 174 IleProAsnIleIleGlyTyrLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyr 193
DB 568 ATTCAAACATCATTTGTTGGCTTGGCTTGCATCTCTTTGCAAAAGACTCATCGTTCTTTAT 627
QY 194 MetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValProValTyr 213
DB 628 ATGGGACCATGTCCTGAGGGGTTTGGTGTGGTGTGATCTCTTATACGGTGCAGGTTTAC 687
QY 214 IleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSer 233
DB 688 ATAGCAGAAATATACCTCAAAACATGAGAGGTGCTCTTGGCTCAGTGAATCAGTTATCT 747
QY 234 ValThrPheGlyIlePheLeuAlaTyrIleuLeuGlyMetPheIleProThrArgLeuLeu 253
DB 748 GTAACCGTGTGTATATTGTTGGCATATTTGTCGGCATGTTGTTCTCTGGAGCTTCTT 807
QY 254 AlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPheIleProGlu 273

```

```

DB 808 CTGTAAATAGGAATCTTGCCTTGCACTGTGTGATACCTGGCCTATCTTCAATCCAGAA 867
QY 274 SerProArgTyrLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnVal 293
DB 868 TCCCAAGATGCTTGGCAAGATGAACATGATGATGATTTTGAGACTTCTTTTCAAGTT 927
QY 294 LeuArgGlyPheGluThrAspIleThrGluValAsnAspIleLysArgAlaValAla 313
DB 928 CTGAGGGGATTTGAGACTGACATCAGCGCGGAAGTGAATGATATTAAGAGAGCAGTAGCG 987
QY 314 SerSerSerLysArgThrThrIleSerPheGlnGlnLeuAsnGlnLysLysTyrArgThr 333
DB 988 TCAGCAAAACAAAGGACAAACGATCCGTTTTCAAGATTTAAACCAAGAAATACCGCACA 1047
QY 334 ProLeuLeuLeuGlyIleGlyLeuValLeuGlnAsnLeuSerGlyIleAsnGlyVal 353
DB 1048 CCCCTAATACTAGGAATGGCTTACTTGTACTGCAACAGCTAAGTGGGAATCAATGGAATA 1107
QY 354 LeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThr 373
DB 1108 TTGTTTATGAGGTAGCATCTTCAAGCAGCAGGTCTCACAAACAGTACTTGGCTACA 1167
QY 374 CysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArg 393
DB 1168 TGTGCACCTTGGTGTATCTTCTTACAGAGTTCACACCTGGTTATTAGACAGA 1227
QY 394 AlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuAla 413
DB 1228 GTGCGCCAGCAGGATCTTCTTATCATCTTCTGCTGSGATGACTCTAAGCCTCTTGA 1287
QY 414 ValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIle 433
DB 1288 GTTGTCTGTTGATTTTCTCAAGGATAGCATTTCAACAGATTTCTCAGATGTACTAC 1347
QY 434 LeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMet 453
DB 1348 TTAAGTATGATCTCTCTGTTGCTCTTGTGGCTTTTGAATCGCCTTCTCTTCGGTATG 1407
QY 454 GlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGly 473
DB 1408 GTTGCCATTCATGATGATATAATGTCAGAGATCTCTCCGCTTAGTATCAAGAGTCTCG 1467
QY 474 GlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsn 493
DB 1468 GGAAGCTTGGCAGCCTGCCAATCTGCTTACATCTTTGGAATTAACAAATGACAGCAAC 1527
QY 494 LeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValValSerAlaPhe 513
DB 1528 TTGATGCTTAGCTGGAGTCTGGAGGACCTTTGTGCTCTACATGCTGTGAGTCTTTC 1587
QY 514 ThrIleValPheValLeuTrpValProGluThrLysGly***AsnSerArgGlyAsp 533
DB 1588 ACCCTCGTGTCTGTCATCTTTGGGTGGCAGAGACAAAAGGAAG-AACTCTCGAAGAGAT 1646
QY 534 ThrIlePheValSerLeuSerIle 541
DB 1647 ACAATGGTCTCTTCCGCTGAGCCTT 1670

```

RESULT 3

```

US-10-425-114-26596
; Sequence 26596, Application US/10425114
; Publication No. US20040003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114

```



```

FEATURE:
; OTHER INFORMATION: Clone ID: LB3689-227-G11_FLI
US-10-425-114-24968

Alignment Scores:
Pred. No.: 136-170 Length: 2049
Score: 1806.50 Matches: 366
Percent Similarity: 81.04% Conservative: 70
Best Local Similarity: 68.03% Mismatches: 79
Query Match: 65.01% Indels: 26
DB: 16 Gaps: 4

US-10-051-909-36 (1-553) x US-10-425-114-24968 (1-2049)
QY 34 LeuArgThrArgValMetGlyGlySerAsnArgGlyAlaGlyAlaGlyGlu 53
DB 314 CTACGGCATGAGCTTCGGGACCCAGAGAGTGGCGGGAGATG-GGGGAGGAGC-TCC 371
QY 54 SerGlySerAspHisAspGlyValLeuArgProLeuLeuAsnThrGlySerTrpTyr 73
DB 372 TCCGCTCCGAC-----CTCGGAAGCGGCTCATCAACCCGGGAGCTGGTAC 419
QY 74 Arg-----MetSerSerArgGlnSerSerPheAlaProGlyThr 86
DB 420 CGATGCGCGCGGGGTGGGTGATGGCTCGCGGCAATCCAGCCTCATGGAGCGCTTG 479
QY 87 SerSerMetAlaVal----LeuArgGluSerHisValSerAlaPheLeuCysThrLeu 105
DB 480 GGCTCCTCTGGTCTCTCCGGACCGAGTCCGATATCCGCCACGCTCTGCACGC-TATT 539
QY 106 ValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerProThrGlnAsp 125
DB 540 GTCCGCTAGTCCATCCAGTTCGGTTTCACATGCGGCTACTCTCCGCCACGAGGAC 599
QY 126 AlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSer 145
DB 600 GCATCATCTGATCTCGGCCCTCTCCTCTGAGTTCTCCCTCTTCGGTTCATTATCT 659
QY 146 AsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArg 165
DB 660 AATATAGGGCGATGGTAGGCGCATCTCCAGTGGGCACTTCAGAGATATATCGGCCG 719
QY 166 LysGlySerLeuMetIleAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPhe 185
DB 720 AAGGGTCTCTCATGATCGCTGCAATCCAAATATATGGTGGTCTCGGATATCATTC 779
QY 186 AlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIle 205
DB 780 GCAAAAGATTCTCTTTCTTTATGGTCTGCTGCTAGAGGATTTGGAGTCGGTGTA 839
QY 206 IleSerTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAla 225
DB 840 ATATCGTATACGATACCGGTTTATATGCAAAATCGCTCTCAGGATCAGAGGGAGCT 899
QY 226 LeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGly 245
DB 900 CTGGTCTCTCAATCAGCTCTCCGTCACGATGGTATATGCTTCCTACTTGTGTGC 959
QY 246 MetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIle 265
DB 960 ATGTTTCTCCCGGAGAAATCTGCTGTTTATAGGCAATTTTACTTGTTCATCTCGATT 1019
QY 266 ProGlyLeuPhePheIleProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGlu 285
DB 1020 CTGGACTGTTCTTTGTCCTGAATCCCAAGTGGTGGCAAAATGGGAGATGGAG 1079
QY 286 AspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrGluVal 305
DB 1080 GATTTCGAATATTCAATGCAAGTCTCGAGAGGATTTTCAGACAGATATCAGCAGAGTA 1139
QY 306 AsnAspIleLysArgAlaValAlaSerSerLysArgThrThrIleSerPheGlnGlu 325
DB 1140 AATGAATAAGAGATCAATTAGCATCATCGAGAGGAGGACCAACCAATAGGTTCCGTGAT 1199

```

RESULT 6

```

US-10-424-599-132422
; Sequence 132422, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 132422
; LENGTH: 2202
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2202)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:

```

```

QY 326 LeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGlyLeuValLeuGln 345
DB 1200 ATCAAAACAGAGAGATACAGTGTCCCTTGTAGTAAGATCGGTCTCTCTCTCTGAG 1259
QY 346 AsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGly 365
DB 1260 CAGCTAAGTGGTGTCAATGGCATTTCTATTTATGTGCGAGCATCTTCAAGAGCTGCTGGT 1319
QY 366 ValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGly 385
DB 1320 ATTACAAACAGTATCTAGCAACATTTGGTTTAGGGGTGTTCAGGTGATTTCTACTGGA 1379
QY 386 ValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIleSerThrSer 405
DB 1380 GTGACAACCTGGTGTGACTGACAAAGCTGGTGAAGGCTTCTCTCAATTATTTCCACCACA 1439
QY 406 GlyMetThrLeuCysLeuLeuAlaValSerValPhePheLeuLysAspAsnIleSer 425
DB 1440 GGAATGGTCATTACTCTTTGTTTCTGTCTGTCATTTTGTGAAGGACAAATAGCT 1499
QY 426 GlnAspSerAsnSerTyrIleLeuThrMetIleSerLeuValGlyIleValSerPhe 445
DB 1500 GCTGCTTCGCACTATATCTCTGTAATGAGTATGCTTTCACCTGGCTGGATTTGGCATT 1559
QY 446 ValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeu 465
DB 1560 GTGATTGCAATTTCTCTTGGCTGGGAGCGATTCGGTGGATCATTTATGTTCTGAGATCCT 1619
QY 466 ProValSerIleLysSerLeuGlySerIleAlaThrLeuAlaAsnTrpLeuThrSer 485
DB 1620 CTTGTTACATCAAGAGCTTCTGGAAAGTGTGGACCCCTGGGAACCTGGCTGACAGCA 1679
QY 486 PheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGlyIleGlyThrPheLeu 505
DB 1680 TGGGCCATTACAAATGACGCAAGCTGATGTTGAACCTGGAGCAGTGGAGGAACATTTGCT 1739
QY 506 SerTyrMetValValSerAlaPheThrIleValPheValValLeuTrpValProGluThr 525
DB 1740 ATCTACGCGCGTGTCTTACCATGGCCCTCATTTTCGTGTGCTTGTGGTGGCTGGAGACC 1799
QY 526 LysGly***AsnSerArgGlyAspThrIlePheValSerLeuSerIleGlnArgGlnLeu 545
DB 1800 AAGGAAG-AACGCTAGAGGAATCGCCTTCTCATTCGCTGACACGTCGTCATGATCTA 1858
QY 546 -GlnTrp-----LeuProGluCysLeu 552
DB 1859 GGTATGGGAAGCCACACACCATGTTGCTTGGTGTCTGCCGCGTGGCTT 1910

```


OTHER INFORMATION: Clone ID: PAT_MRT3847_90584C.1
US-10-424-599-132422

Alignment Scores:

Pred. No.: 1.07e-161 Length: 2202
Score: 1718.00 Matches: 334
Percent Similarity: 83.67% Conservative: 76
Best Local Similarity: 68.16% Mismatches: 76
Query Match: 61.82% Indels: 5
DB: 16 Gaps: 2

US-10-051-909-36 (1-553) x US-10-424-599-132422 (1-2202)

```
QY 52 GluCluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAnThrGlySer 71
DB 255 GAAGAGAGTGGCGATGCGAGGGAT-----CTTCAGAAACCGTTCCTCCACAGGGAGT 308
QY 72 TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal 91
DB 309 TGGTACAAGATGGGTTCAGGCAGTCCAGCATCATG-----GGATCCTCCACTCATGTT 362
QY 92 LeuArgGluSerHisValSerAlaPheLeuCysThrLeuLeuValAlaLeuGlyProIle 111
DB 363 ATCCGCGACGGCGCGTTCCTGCTCTTCTGCGTCTCATCGTTCGCTTGGGTGCCAATT 422
QY 112 GlnPheGlyPheThrSerGlyPheSerProThrGlnAspAlaValMetValArgAspLeu 131
DB 423 CAATCGCTTCACGTGGGTATCTCTCCACCCAGGGGCTATAGTTCGCGATCTA 482
QY 132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyMetVal 151
DB 483 AACCTCTATTTCCGAGTTTCTTCTTTGGATCTTTGTCTAATGTTGGGAGCAATGGTG 542
QY 152 GlyAlaIleAlaSerGlnMetAlaGluTyrIleGlyArgGlyGlySerLeuMetIle 171
DB 543 GGAGCTATAGTGTGTCAGATGCTGAATACATCGCGCGCAAGGGTCATTGTGATGATT 602
QY 172 AlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPhe 191
DB 603 GCTCGCATCCCAATATAATAGGGTGGCTTGTCTATTTCTTTGCCAAGATTCTCGTTT 662
QY 192 LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleSerTyrThrValPro 211
DB 663 TTGTATATGGGAGGTGTTTGGAGGTTTGGCGTGGGATATCTCTTATGTGGTGCCT 722
QY 212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231
DB 723 GTTTATATAGCTGAGATTGCACCTCAAAACTTGAGAGTGGCCTTGGTCACTGAACCCAG 782
QY 232 LeuSerValThrPheGlyIlePheLeuAlaIleLeuLeuGlyMetPheIleProTrpArg 251
DB 783 CTCCTCTGTACAATTGGCAATTATGCTGGCTTATCTGTGGTCTTTTGTCAACTGGAGA 842
QY 252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPheIle 271
DB 843 GTCCTTCGAATTTAGGAATTTTCCCTGTGACAGTATTAATACCTGGATATTTTTCATA 902
QY 272 ProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu 291
DB 903 CCTGAATCCCCACAGTGGTGGCCACAGTGGGATGATAGATGAGTTTGAGACTTCTTTG 962
QY 292 GlnValLeuArgGlyPheGluThrAspIleThrGluValAsnAspIleLysArgAla 311
DB 963 CAAGTGTGTACGAGATTTCACATGATATATCTGTTGAAGTACATGAATTAAGAGATCT 1022
QY 312 ValAlaSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyr 331
DB 1023 GTGGCTTCAACGGGAAAAGAGCTGCGAATTCGATTCAGATCTCAAGAGAAAAGATAT 1082
QY 332 ArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsn 351
DB 1083 TGGTTCCCGTTAATGTGGTATTTGGATTACTTGTCTTCAGCAATTTATCTGGTATCAAT 1142
```

```
QY 352 GlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeu 371
DB 1143 CGAATTTTGTCTATTCAACTACCATCTTGCAAATGCGAGGAATTCATCCAGCGAAGCT 1202
QY 372 AlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeu 391
DB 1203 GCTACAGTTGGACTTGGAGCGCTTCAGGTTCATAGCACTGGAATTCACATGTTGGTG 1262
QY 392 AspArgAlaGlyArgArgIleLeuLeuIleLeuSerThrSerGlyMetThrLeuCysLeu 411
DB 1263 GACAAAAGTGGCGGAGGCTGCTCTTAATAATATATCTCATCTGTATGACAGTTAGCCTT 1322
QY 412 LeuAlaValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyr 431
DB 1323 CTCATGTTTCTATAGCAATTTATCTGGAGGGTGTGTATCAGAGATTCACATTTATTC 1382
QY 432 TyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPhe 451
DB 1383 AGCAATTTGGGAATAGTTTCTATTCTTGGACTCGTGGCTATGGTGTGTTCTCTCTA 1442
QY 452 GlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLysSer 471
DB 1443 GGTCTGGGACCCATCCCTTGGCTTATAATGCTCAGATACCTTCCAGTGAATATAAGGC 1502
QY 472 LeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThr 491
DB 1503 CTTGCTGGCAGCATAGCGACAATGGAAATTTGGCTGATTTCGTGGGGATCACGATGACT 1562
QY 492 ThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValValSer 511
DB 1563 GTAACTTCTGCTTTTGAATTTGGAGCAGTGGAGGACATTTACAATCTACAGTCTGATCT 1622
QY 512 AlaPheThrIleValPheValLeuTrpValProGluThrLysGly***AsnSerArg 531
DB 1623 GCCTTTACTATAGCTTTTATAGCAATGTGGGTCTCTGAGACCAAGGAGAACATTGGA 1681
QY 532 GlyAspThrIlePheValSerLeuSerIle 541
DB 1682 AGAAATTCAGTTTCTCTTCAGATAGATATA 1711
```

RESULT 7

US-10-437-963-63465
Sequence 63465, Application US/10437963
Publication NO. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 63465
LENGTH: 1304
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_64704C.1

US-10-437-963-63465

Alignment Scores:
Pred. No.: 1.45e-127 Length: 1304
Score: 1374.00 Matches: 233
Percent Similarity: 68.89% Conservative: 48
Best Local Similarity: 59.19% Mismatches: 58
Query Match: 49.44% Indels: 97


```
QY 53 userGlySerAspHisaspGlyValLeuArgArgProLeuLeu---AsnThrGlySerTr 72
Db 122 GAGGGGAGGACACGACGCGCGGCTGCGAGAGCGGCTGCTGCCAACAGCGGAGCTG 181
QY 72 pTyArg-----MetSerSerArgGlnSerSerPhe---AlaProGlyThrSerSerMe 89
Db 182 GTACCGGATGGGATGGGTCGCGCCAGTCCAGCTCAAGCGCCGCGGACCTCCTCCAT 241
QY 89 tAlaValLeuArgGluSerHisValSerAlaPheLeuCysThrLeuLeuValAlaLeuG 109
Db 242 GCGCGTCTCGCGGAGTCCGACGCTCGCCCTCTCTGACGCTCATCGCGGCTCGG 301
QY 109 yProiledGlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValAr 129
Db 302 CCCCATCCAGTTCGCTTCCCGCGGCTACTCTCTCCCGAGCGAGCCAGCATCATCG 361
QY 129 gAspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyG 149
Db 362 GGACTTCAACCTCTCCATCTCCGAGTCTCGGTTCGCGTCTCGTGCCTCAACGTCGGCGC 421
QY 149 yMetValGlyAlaIleAlaSerGlyGlnMetAlaGlyTyIleGlyArgGlySerLe 169
Db 422 CATGTCGGGCGATCGCCAGCGCCAGATGCGCGAGTACGTTGCGCGCAAGGTCGTT 481
QY 169 uMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAl 189
Db 482 GATGATTGCAGCAATTCGAAACGTCATTGGTTGGCTTCGCGATCTCTTTCGAAAGACTC 541
QY 189 aserPheLeuTyMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyTh 209
Db 542 TTCATTCTGTATATGAGACCTTGCTTGAAGATTGCGTGTGCGTATCATTTCTTCTATGT 601
QY 209 rValProValTyIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerVa 229
Db 602 GGTACCGGTATACATAGCAGAGATATCTCCACAAACATGAGAGGGGCTCTTGCGCTCTGT 661
QY 229 lAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyIleLeuLeuGlyMetPheIlePr 249
Db 662 GAACGAGTATCTGTAAACCTTGTGTATCATGTTTGCCTACTTGTCTCGGCTGTGTGTTC 721
QY 249 oTriArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPh 269
Db 722 TTGAGGGCTCTTCGAGTAATAGAACCTTGCCCTGCTATGATGTGTGATCTGGCCCTTTT 781
QY 269 ePheIleProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluTh 289
Db 782 CTTTCATTCGGAATCTCCAGATGGCTGCGCAAGATGAATATGATGATGATGATGATGATG 841
QY 289 rSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGluValAlaAsnAspIleLy 309
Db 842 TTCTCTACAGTCTTGAGAGGATTCGATGCTGATCATCTGCGAGAAATGATGATATATAA 901
QY 309 sArgAlaValAlaSerSerLysArgThrThrIleSerPheGlnGluLeuAsn 327
Db 902 GAGAGCAGTAATGTGAGCAAAACAAAGGGCTACAAATTCGTTTCCAGAGTTGAAC 956
```

RESULT 9

```
US-10-424-599-131066
; Sequence 131066, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 131066
```

```
; LENGTH: 1283
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_9935C.1
; US-10-424-599-131066

Alignment Scores:
Pred. No.: 1,23e-108 Length: 1283
Score: 1185.00 Matches: 223
Percent Similarity: 81.84% Conservative: 61
Best Local Similarity: 64.27% Mismatches: 51
Query Match: 42.64% Indels: 12
DB: 16 Gaps: 2

US-10-051-909-36 (1-553) x US-10-424-599-131066 (1-1283)
```

```
QY 201 PheGlyValGlyIleIleSerTyThrValProValTyIleAlaGluIleSerProGln 220
Db 2 TTGGGGGTAGGAATAATATCTTACACGGTGCCTGTGTACATAGCTGAGATATCACCTCCA 61
QY 221 AsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeu 240
Db 62 AACTTGAGGGGGGTTTGGTTTCAGTTAACCCAGCTCTCTGCCATTGGAATTATGCTG 121
QY 241 AlaTyIleLeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuPro 260
Db 122 GCATATCTCTCGGGGATTTTGTGAATGGAGAAATCTTGCATTCGAGTCCCTAGATGGTGGCAAAA 241
QY 261 CysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeuAlaLys 280
Db 182 TGTCATATATGATACCTGGCTATTTTCATTCGAGTCCCTAGATGGTGGCAAAA 241
QY 281 MetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAsp 300
Db 242 ATGGGAATGACAGAGAAATTTGAAACTTCCTTGCAGTCTCGAGGCTTTCGAGCTGAT 301
QY 301 IleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerLysArgThrThr 320
Db 302 ATTCTGTGTGAAGTGAATGAATTAAGAGGCTGTGTCTTCAACAAAGAGAACTACA 361
QY 321 IleSerPheGlnGluLeuAsnGlnLysLysTyArgThrProLeuLeuLeuGlyIleGly 340
Db 362 GTTCGATTTGCAGACCTCAACAAAGAGATATTTGGCTTCCCTTAATGATTGGAATTGCA 421
QY 341 LeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyThrAlaSerSerIle 360
Db 422 CTGCTTATTTTTCACACAGCTTTCTCGAATTAATGGTGTCTTTTATTTCCAGTACCATC 481
QY 361 PheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGln 380
Db 482 TTTCGAAATGCTGGAATCAGTTCAAGTACGCGGCAACATTCGGAGTTGGTCTGTTTCA 541
QY 381 ValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeu 400
Db 542 GTTCTTGCCACCACTTAATCTTGTGGCTGGCAGACAAATCTGTGCGCGGCTTCTCCTT 601
QY 401 IleIleSerThrSerClyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeu 420
Db 602 ATTTGCTCTGCAACTGGAATGCTTTTAGTCTCTTAGTTGTTGCAATTCATTTACATPA 661
QY 421 LysAspAsnIleSerGlnAspSerAsnSerTyTyIleLeuThrMetIleSerLeuVal 440
Db 662 AAGGCTAGTATATCAGAAACTTCTTCTTGTATGGGATATTCAGCACCTTGTGCTGTT 721
QY 441 GlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleLeuProTrpLeuMet 460
Db 722 GGAGTCGTGGCATGGTATTATTCATTTCTCTTGAATGGAGCTATGCGATGGATTATA 781
QY 461 MetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAla 480
Db 782 ATGTCGTGAGATCTTCCAAATAACATTAAAGGCTTCGCGGAGAGTGTGGCAACACTTGC 841
```



```

; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_90588C.1
US-10-424-599-132426

```

Alignment Scores:

Pred. No.:	9, 29e-104	Length:	2105
Score:	1139.50	Matches:	228
Percent Similarity:	70.55%	Conservative:	69
Best Local Similarity:	54.16%	Mismatches:	58
Query Match:	41.00%	Indels:	67
DB:	16	Gaps:	2

US-10-051-909-36 (1-553) x US-10-424-599-132426 (1-2105)

	QY	187	LysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIle	206
	DQ	320	CAGGAATTCCTCATTTTGTATATGGGAGCCTGCTGAAAGTTTTGGCGFCGGGATTACG	379
	QY	207	SerTyThrValProValTyrlleAalagIulleSerProGlnAsnMetArzGlyAlaLeu	226
	DB	380	TCTTATTAGGTGCCTGTTTATATAGCTGAGATTGCACCTCAAAACTTGAGAGGTGCCCTT	439
	QY	227	GlySerValAsnGlnLeuSerValThrPheGlyIlePheIleuAlaTyrrLeuLeuGlyMet	246
	DB	440	GGATCAGTGAACACAGCTCTTATTACTATTGGCCATTATGCTGGCTTATCTTTTGGGCCCTT	499
	QY	247	PheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIlePro	266
	DB	500	TTTGTCAACTGGAGAGTGCITGCAATCTAGGAATTTTGCCCTGTACAGTAATTAATACCT	559
	QY	267	GlyLeuPhePheIleProGluSerProArgTrpLeu	278
	DB	560	GGATTATTTTTCATACCTGAATCCCCAGATGGTGTGATGGATATGTTATGCAATTATAT	619
	QY	278	-----	278
	DB	620	TGATCAAAGTTTTGTACTTTTCATCTTTGTAAAATACTTGTTCCTTTCAGAAAAAGGATG	679
	QY	278	-----	278
	DB	680	GTAACCTTTTCTACTCTTGGTAATTATCTTTCTGCTCTCTTTCAATTTATTTTATTTTA	739
	QY	279	-----AlaLysMetAsnLeuthrGluAspCy	287
	DB	740	TTTTCTGGTGGTGGGAAATCGGGGAAGTGTAGCTTAAGATGGGATGACAGATGAGTT	799
	QY	287	sGluThrSerLeuGlnValLeuArgGlyPheGluThrAspilethrThrrGluValAsnAs	307
	DB	800	TCGAGACTCTCTTTGCAAGTGTTACGAGGATTTTGACACAGATATATCTGTTGAAGTATATGA	859
	QY	307	pIleLysArgAlaValAlaSerSerLysArgThrThrIleSerPheGIndgIuLeuAs	327
	DB	860	AATTAAAGATCTGTGGCTTTCAACGGGAAAAAGAGCTACAAATCGATTCGATATCTCAA	919
	QY	327	nGlnLysTyrrArgThrProLeuLeu-----LeudglyIleGlyLe	341
	DB	920	GAGGAAAAAGATATTGGTTCCTCCCTTAATGSTATATAATACACATGTTAGTTGGTATGGGATT	979
	QY	341	uLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrrAlaSerIlePh	361
	DB	980	ACTTGTCTTCAGCAGGTATCTGGTATCAATGGAGTTTTGTTCTATTCAACTACCATCTT	103
	QY	361	eLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnVa	381
	DB	1040	TGCAAAATGCAGGAATTTGGTCAGCGAAGCTGCTACTGTTGGACTTGGTCCCGCTTCAGGT	109
	QY	381	IleuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyVargArgIleLeuLeuIl	401
	DB	1100	CATAGCACTGGAAATTCACGTGGTGGTGGACAAGAAGTGGTCGGAGGCTGCTTCTAAT	115
	QY	401	erLeSerThrSeryMetThrLeuCysLeuLeuAlaValSerValValPhePheLeuTy	421

1160	GATATCCTCATCTGTAATGACTGTTAGCCTTCTCATGTTTCAATAGCATTTTATCTGGA	1219
421	sAspAsnIleSerGlnAspSerAsnSerTyrTrpIleLeuThrMetIleSerLeuValGl	441
1220	GGGGTGTGATCTCAGGATTCACATCTATTTCAGCATGTTGGGAATAGTTCTTGTTGTGG	1279
441	YIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMe	461
1280	ACTGTGGTTATGGTCATTGGGTTCTCTCTGTGCTGGGACCATCCCTTGGGTTATAAT	1339
461	tSerGluIleLeuProValSerIleIleYsSerLeuGlyGlySerIleAlaThrLeuAlaAs	481
1340	GTCTGAGACTTCCAGTGAAATATAAGGCCCTTGTGGCAGCATACGGACATGGGAAA	1399
481	nTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGl	501
1400	TTGGCTGATTTCGTGGGTGATCAGCATGACCGCTAACTTACTTTGAAATGGGAACAGTGG	1459
501	YGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValValLeuTr	521
1460	AGGGACATTTACAATCTTACACAGTGTAGTGCCTTACTATTGCTTTTATACGATTATG	1519
521	pValProGluThrLysGly***AsnSerArgGlyAspThrIlePheValSerLeuSerIle	541
1520	GGTTCCTGAGACCAAGGAAG-AACATTGGAAAGAAATTCAGTTTCTTCAGATAGATGT	1578
541	e 541	
1579	T 1579	

RESULT 12

```

US-10-425-114-5566
; Sequence 5566, Application US/10425114
; Publication No. US20040034888A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
;
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
;
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
;
; NUMBER OF SEQ ID NOS: 73128
;
; SEQ ID NO 5566
;
; LENGTH: 1027
;
; TYPE: DNA
; ORGANISM: Zea mays
;
; FEATURE:
;
; OTHER INFORMATION: Clone ID: 700464574_FLI
;
; US-10-425-114-5566

```

Alignment Scores:

Argument Scores:					
Pred. No.:	3.58e-104	Length:	1027		
Score:	1139.00	Matches:	233		
Percent Similarity:	99.57%	Conservative:	0		
Best Local Similarity:	99.57%	Mismatches:	1		
Query Match:	40.99%	Indels:	1		
DB:	16	Gaps:	0		

US-10-051-909-36 (1-553) X US-10-425-114-5566 (1-1027)

Qy	320	ThrIleSerPheGlnGlnLeuAsnGlnIleLysIleValTyrArgThrProLeuLeuLeuGlyIle	339
Db	3	ACAATCAGTTTTCAAGATTAAACCAAAAGAAATACCGCAGCCACTTACTTCAGGGATT	62
Qy	340	GlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaLeuSerSer	359
Db	63	GGCCCTACTTGTTACTGCCAAATCTTACTGGGAATCAACGGTGTTACTGTTTATGCAAGTAGC	122

QY 360 IlePheIysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaAla 379
 Db 123 ATCTTCAAGAGTGAGGGGTTACAAACAGGACCTGGCCACCTGTTCACTTGGTGTATT 182
 QY 380 GlnValLeuAlaThrGlyValThrThrTriLeuLeuAspArgAlaGlyArgAlaLeu 399
 Db 183 CAGTCTCTGCTACTGAGATTACAAATGGCTGTAGACGAGCTGGACGACGATCCCT 242
 QY 400 LeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValPhePhe 419
 Db 243 CTCATTATTTCTACCTCTGGCATCACTCTATGCCTTCTTGGCGTTCTGTGTATTTT 302
 QY 420 LeuIysAspAsnIleSerGlnAspSerAsnSerTyrrIleLeuThrMetIleSerLeu 439
 Db 303 CTCAGGATAACATTTTACAGGATTTCACTCATACATCTTAAACATGATCTCCCT 362
 QY 440 ValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeu 459
 Db 363 GTTGGTATTGTCTTTTGTTCATTACCTTCTCGTTTGGTATGGTGCCATTCCATGGCTC 422
 QY 460 MetMetSerGluIleLeuProValSerIleIysSerLeuGlyGlySerIleAlaThrLeu 479
 Db 423 ATGATGTCTGAGATCTCCCGGTTAGCATCAAGAGCCTTGGCGAAGCATCGCAACACTG 482
 QY 480 AlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSer 499
 Db 483 GCCAACTGGCTGACATCTTTCGCCATTAACATGACGACGAACTTGATGCTCACGTGGAGT 542
 QY 500 ValGlyGlyThrPheLeuSerTyrrMetValValSerAlaPheThrIleValPheValVal 519
 Db 543 GTTGGAGGCACTTTCTCTCTACATGTTGTGAGCGCCTTCCACCATCGTTTGTGTGTC 602
 QY 520 LeuTrpValProGluThrLysGly**AsnSerArgGlyAspThrIlePheValSerLeu 539
 Db 603 CTTTGGGTCGGGAGACGAGGGAG-AACTCTTAGAGAGATACAAATTTTCGTTTCGCTG 661
 QY 540 SerIleGlnArgGlnLeuGlnTrpLeuProGluCysLeuSer 553
 Db 662 AGCATTACAGTCACTGCAATGTTGTCGCCAGTGTATCT 703

RESULT 13

US-10-767-795-566
 ; Sequence 566, Application US/10767795
 ; Publication No. US20040181830A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Zhou, Yinhua
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; Publication of Invention: Plants and Uses Thereof For Plant Improvement
 ; FILE REFERENCE: 38-21(53534)B
 ; CURRENT APPLICATION NUMBER: US/10767,795
 ; CURRENT FILING DATE: 2004-01-30
 ; NUMBER OF SEQ ID NOS: 117596
 ; SEQ ID NO 566
 ; LENGTH: 990
 ; TYPE: DNA
 ; ORGANISM: Gossypium hirsutum
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C14353_1

US-10-767-795-566
 Alignment Scores:
 Pred. No.: 6.61e-85 Length: 990
 Score: 946.50 Matches: 196
 Percent Similarity: 80.55% Conservative: 40
 Best Local Similarity: 66.89% Mismatches: 53
 Query Match: 34.06% Indels: 5
 Ds: 17 Gaps: 1
 US-10-051-909-36 (1-553) x US-10-767-795-566 (1-990)
 QY 53 GluSerGlySerAspHisAspGlyValLeuArgProLeuLeuAsnThrGlySerTrp 72

Db 124 GATGATGCTGAAGTGAAGAGGGAAGTCTAAGAGGACCATCTCTACATCTGGAAGCTGG 183
 QY 73 TyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaValLeu 92
 Db 184 TATCGTATGGGTCTTAGGATGGGGTCTAGTATGTAGGATCTCTCAA-----GCTCTT 237
 QY 93 ArgGluSerHisValSerAlaPheLeuValLeuLeuIleValAlaLeuGlyProIleGln 112
 Db 238 CGTGATAAATCACTCGGTTGAGCTTGTGTTTGTATGTTGTTGGTCCCATCCAA 297
 QY 113 PheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeuAsn 132
 Db 298 TTTGGTTTCACTTCTGGTTACTTCTCCGACACAATCCGCAATCATCGAGATCTTGGGA 357
 QY 133 LeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGly 152
 Db 358 CTAACAGTCTCAGAGTTTCTTCTATTTGGTCTTTATCAATGTTGGTGCCATGTTGGA 417
 QY 153 AlaIleAlaSerGlyGlnMetAlaGluTyrrIleGlyArgLysGlySerLeuMetIleAla 172
 Db 418 GCATAGCCAGTGGTCAGATAGCCGAAATGATTTGACCCGAAAGGCTCTTTAATGACTGCT 477
 QY 173 AlaIle-ProAsnIleIleGlyTrpLeuAlaIleSerPheAlaIysAspAlaSerPheLe 192
 Db 478 GCAATTCCTTAATAATTTGGATGGCTTGTATATCTTTTCAAGAGATCTCTCTTTTCT 537
 QY 192 TyrrMetGlyArgLeuLeuGlyPheGlyValGlyIleIleSerTyrrValProVa 212
 Db 538 TTACATGGAGAGCTGTTGCAAGGTTCTCGTGTCCGATATCTCTTATACGGTGCCTGT 597
 QY 212 TyrrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLe 232
 Db 598 ATATATAGCTGAGATAGCACCTCAGAAATTTGAGGGGAGCTTGGGTTTCAGCGAATCAGCT 657
 QY 232 uSerValThrPheGlyIlePheLeuAlaTyrrLeuLeuGlyMetPheIleProTrpAcgLe 252
 Db 658 ATCTGTCACTCTCGGAATAATGCTGGCCTATCTACTTCGACTTTTGTGTAGTCCCGAT 717
 QY 252 uLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIlePr 272
 Db 718 ACTCGCGCTTTTAGGAATACCTCTTGTACAATTTGATACCTGCTCTATTTTTCATTC 777
 QY 272 oGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuG 292
 Db 778 AGAATCTCTCGATGGCTGGCAAACTGGGTATGACAGAGATTTTCGAAATCTCTTTTGA 837
 QY 292 nValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysAAlaVa 312
 Db 838 AGTTCTCAGGTGCTCCGATGCTGATATTTCTATTGAGTGAATGAATCAACAGATCTGT 897
 QY 312 lAlaSerSerSerLys-ArgThrThrIleSerPheGlnGluLeuAsnGlnIlysIysTyra 332
 Db 898 CTCATCAACCACTAGCAAGAACTACGATTTCGTTG--CACATCTCAAGAAAGATATT 955
 QY 332 rgThrProLeuLeuLeuGlyIleGlyLeuLeuVal 343
 Db 956 GGTCTCCATTGATGTTGGAAATTCGCTCACTATG 990

RESULT 14

US-09-938-842A-2331
 ; Sequence 2331, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SCRI1300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24

;; PRIOR APPLICATION NUMBER: US 60/227,866
;; PRIOR FILING DATE: 2000-08-24
;; PRIOR APPLICATION NUMBER: US 60/264,647
;; PRIOR FILING DATE: 2001-01-16
;; PRIOR APPLICATION NUMBER: US 60/300,111
;; PRIOR FILING DATE: 2001-06-22
;; NUMBER OF SEQ ID NOS: 5379
;; SEQ ID NO 2331
;; LENGTH: 1473
;; TYPE: DNA
;; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2331

Alignment Scores:
Pred. No.: 4,81e-82 Length: 1473
Score: 920.50 Matches: 186
Percent Similarity: 61.69% Conservatives: 91
Best Local Similarity: 41.43% Mismatches: 159
Query Match: 33.12% Indels: 13
DB: 9 Gaps: 3

US-10-051-909-36 (1-553) x US-09-938-842A-2331 (1-1473)

Qy 98 SerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSer 117
Db 151 TCTGTTTCTCAGACCTTTGTCGGTATACAGGCTCTCTCTACCGGTTGTGGCGTT 210
Qy 118 GlyPheSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGlu 137
Db 211 GGTITTTTCACCGGTGCACAGCGGATTTACCAAGATTATCTCTCCGTTCAGAA 270
Qy 138 PheSerAlaPheGlySerLeuSerAsnValGlyMetValGlyAlaIleAlaSerGly 157
Db 271 TACTCAATGTTGGGTGCATCTTGACATTAGGAGGCTTGATCGGTGAGTATTCAGCGGT 330
Qy 158 GlnMetAlaGluTyrIleGlyArgGlySerLeuMetIleAlaIleProAsnIle 177
Db 331 AAGTCGCTGATGCTTGGGAAGAAACCGGACGATGTTGTTTCGGAATTTCTCTGTATC 390
Qy 178 IleGlyTyrLeuAlaIleSerPheAlaIleAspAlaSerPheLeuTyrMetGlyArgLeu 197
Db 391 ACAGCTGGCTTCTGTAGCATTTGGCTCAGATGTCATGTCAGTTCGGAATTCGGA 450
Qy 198 LeuGluGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIle 217
Db 451 TTACTTGAATCGCGGTGTGTATATTTAGTACGTCGATCGGTGTATATAGCGGAAT 510
Qy 218 SerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGly 237
Db 511 GCACCTAAACATGTCGAGGATCGTTGTTGTCGCCAATCAGTTGATGCAAAATTCGGA 570
Qy 238 IlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaIleGly 257
Db 571 ATTTCACTCTCTTCATCATTTGGCAATTTTATTCATGGAGACTACTAACAGTAGTCGA 630
Qy 258 AlaLeuProCysThrMetLeuIleProGlyLeuPheIleProGluSerProArgTyr 277
Db 631 TTGGTGCATGTGTGTCACGCTCTTTGTTTATTTTTCATCCCGAATCTCCAAAGATG 690
Qy 278 LeuAlaIleMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPhe 297
Db 691 CTGGCGAAGTATAGTCGTGATAAGATCGGATCTTCGTTGCAACGCTTAGGGATCT 750
Qy 298 GluThrAspIleThrThrGluValAsnAspIleIleArgAlaValAlaSerSerLys 317
Db 751 GACGTCGATATTTCTGTAAGCAACACCAATTAGAGATACCATTTGACATGACAGAAAC 810
Qy 318 ArgThrThrIleSerPheGlnGluLeuAsnGlnIleLysTyrArgThrProLeuLeuLeu 337
Db 811 GGTGGTGAACATAGATGCTGAATTTGTTTCAGAGACGATACGATATCGTTAATATC 870
Qy 338 GlyIleGlyLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAla 357

Db 871 GGAGTTGGTTTAAATGTTTTCGCAACAATTGTGTGGAGCTCGGTTGTACCTATTATGCT 930
Qy 358 SerSerIlePheLeuAlaIleGlyValThrAsnSerAspLeuAlaIleThrCysSerLeuGly 377
Db 931 AGTAGCTCTCTTCACAAAGAGGATTTCCA---AGTGTATTGGCACATCCGTAATAGCC 987
Qy 378 AlaIleGlnValLeuAlaIleThrGlyValThrTrpLeuLeuAspArgAlaGlyArg 397
Db 988 ACAATTATGTTCCAAAGCAATGCTGGCAACAGTCTTAGTCGATAAAATGGGAGGAGA 1047
Qy 398 IleLeuLeuIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValVal 417
Db 1048 ACGCTCCTAATAGCTTCTTGTTCGAAATGGGTTTGAAGTCTTCTTAAGTGT--- 1104
Qy 418 PhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeu----- 434
Db 1105 -----TCTTACGGTTTCCAGTGTGTTGGCATTTCTCCAGAACTC 1143
Qy 435 ---ThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMet 453
Db 1144 ACTCCCATCTTCACTTGCATCGGCTCTTGGTGCACATTTGTGTCAATTTGCCATGGGAATG 1203
Qy 454 GlyAlaIleProTyrLeuMetMetSerGluIleLeuProValSerIleIleSerLeuGly 473
Db 1204 GGAGGACTACCATGATTAATGCTGAGATATTTCCGATGAATGTGAAGTGTCACT 1263
Qy 474 GlySerIleAlaIleThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrAsn 493
Db 1264 GGGACCTTAGTTACTGTAAACCAATGTTTATTGTTGGATTATCACAATCACTTTCAAT 1323
Qy 494 LeuMetLeuThrTrpSerValGlyThrPheLeuSerTyrMetValValSerAlaPhe 513
Db 1324 TTTATGCTAGAAATGGAATGCAATGCAATGTTTCTTCATCTTCAATGGTCTCCCGCATG 1383
Qy 514 ThrIleValPheValValLeuTrpValProGluThrLysGly**AsnSerArgGlyAsp 533
Db 1384 TCGATCGTATTTATATACATTTTGTGCTACTGACAGCAAAAGCCGATCACTTGAAGAAATA 1443
Qy 534 ThrIlePheValSerLeuSerIleGln 542
Db 1444 CAAGCACTGCTCAACAACATCTGTGCAA 1470

RESULT 15

US-09-938-842A-2331
; Sequence 2331, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2331
; LENGTH: 1473
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2331

Alignment Scores:
Pred. No.: 4,81e-82 Length: 1473
Score: 920.50 Matches: 186
Percent Similarity: 61.69% Conservatives: 91

Best Local Similarity: 41.43% Mismatches: 159
Query Match: 33.12% Indels: 13
DB: 11 Gaps: 3

US-10-051-909-36 (1-553) x US-09-938-842A-2331 (1-1473)

Qy 98 SerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSer 117
Db 151 TCTGTTTTCCTCAGCAGCTTTGTCGGTATCAGGCTCCTCTGTACCGGTTGTGGCGTT 210
Qy 118 GlyPheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGlu 137
Db 211 GGTTTTTCATCGGTGCACAGCAGGAGTACCAAGATTTATCTCTCTCCGTTCAGAA 270
Qy 138 PheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGly 157
Db 271 TACTCAATGTTCCGGTTCGATCTTGACATTAGGAGGCTTGATCGGTGCAGTATTCACGGGT 330
Qy 158 GlnMetAlaGluTyrIleGlyArgIleGlySerLeuMetIleAlaAlaIleProAsnIle 177
Db 331 AAGTCGCTGATGCTTGGGAGAAACGACGATGTTGTTTTCGGAATTCCTCTGATC 390
Qy 178 IleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeu 197
Db 391 ACAGGCTGGCTTTGTGTAGCATTTGGCTCAGAAATGCTGAGTGGCTGGAGATTG 450
Qy 198 LeuGluGlyPheGlyValGlyIleSerTyrThrValProValTyrIleAlaGluIle 217
Db 451 TTACTTGAATCGCGTGGTGGTATTTAGTACGTGATTCGGTGTATATAGCCGAATTT 510
Qy 218 SerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGly 237
Db 511 GCACCTAAACATGCCAGGATGCTTGTGTGCCAATCAGTTGATGCCAAATTCGCGA 570
Qy 238 IlePheLeuAlaTyrLeuLeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGly 257
Db 571 ATTCACCTCTTCATCATTTGGCAATTTTATTCATGGAGACTACTAACAGTAGTCGGA 630
Qy 258 AlaLeuProCysThrMetLeuIleProGlyLeuPheIleProGluSerProArgTyr 277
Db 631 TTGGTGCAATGTTGCTCCAGCTCTTTGTTTATTTTTCATCCCGAATTCACAGATGG 690
Qy 278 LeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPhe 297
Db 691 CTGGCGAAGTTAGGTCGTGTGAAGAATGCCGATCTTCGTTCGAACGCCCTTAGGGGATCT 750
Qy 298 GluThrAspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerLys 317
Db 751 GACGTGCATATTTCTCGTAAGCAACACATATAGATACCATTTGACATGCAGAAAC 810
Qy 318 ArgThrThrIleSerPheGlnGluLeuAsnGlnLysTyrArgThrProLeuLeuLeu 337
Db 811 GGTGGTGAACCTAAGATCTCTGAATTTCTTTCAGACAGATACGCATATCCGTTAATTATC 870
Qy 338 GlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAla 357
Db 871 GGAGTTGTTTAAATGTTTTTTCGAACAAATGTGTGGGAGCTCCGGTGTACCTATTATGCT 930
Qy 358 SerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGly 377
Db 931 AGTAGCTCTTCAACAAGGAGGATTTCCA--AGTGTATTGGCACATCCGTAATAGCC 987
Qy 378 AlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArg 397
Db 988 ACAATATGTTTCCAAAGCAATCTCGCAACAGCTCTCGATGATAAATGGGGAGAGA 1047
Qy 398 IleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValVal 417
Db 1048 ACGTCTCTAATGGCTTCTTGTCTCGCAATGGGTTTGAGTCTTGTCTTAAAGTGT--- 1104
Qy 418 PhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeu----- 434
Db 1105 -----TCTACGGTTTCAGTCTCGGTTGGCATTTCTCCAGAACTC 1143

Qy 435 ---ThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMet 453
Db 1144 ACTCCCATCTTCACATTCGGGTCTTGGGTGCATATTGTGTCTTCCATGGGAATG 1203
Qy 454 GlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGly 473
Db 1204 CGAGCACTACCATCGATTTATATGGCTGAGATATTTCGATGAATGTGAAAGTGTCACT 1263
Qy 474 GlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrAsn 493
Db 1264 GGGACCTTAGTTACTGTAAACCAATGGTTATTTGGTTGGATTATCACAACACTTCAAT 1323
Qy 494 LeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValValSerAlaPhe 513
Db 1324 TTTATGCTAGATGGATGGATGCATCAGGAATGTTCTCATCTTCTCAATGTTCTCGGCCAGT 1383
Qy 514 ThrIleValPheValValLeuTrpValProGluThrLysGly***AsnSerArgGlyAsp 533
Db 1384 TCGATCGTATTATATATACTTTTGGTACCTGAGACAAAGCCGATCATCTTGAAGAAATA 1443
Qy 534 ThrIlePheValSerLeuSerIleGln 542
Db 1444 CAAGCACTGCTCAACAACACTCTGTGCA 1470

Search completed: October 14, 2004, 00:03:56
Job time : 825.832 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 11:33:22 ; Search time 14.714 Seconds
(without alignments)
3616.147 Million cell updates/sec

Title: US-10-051-909-36

Perfect score: 2779
Sequence: 1 PSSSSPRPAGKKKKKNQG.....TIFVSIGRQLWLPECLIS 553

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1757.5	63.2	490	2 T14545	probable sugar tra
2	1734.5	62.4	487	2 E56782	hypothetical prote
3	920.5	33.1	496	2 T52132	probable sugar tra
4	824	29.7	483	2 D36589	hypothetical prote
5	820.5	29.5	474	2 B86221	hypothetical prote
6	808	29.1	454	2 A86221	hypothetical prote
7	748.5	26.9	461	2 G85059	probable sugar tra
8	729	26.2	457	2 H85059	probable sugar tra
9	671.5	24.2	378	2 G86220	hypothetical prote
10	626	22.5	490	2 F86220	hypothetical prote
11	621	22.3	348	2 D84922	probable sugar tra
12	578.5	20.8	521	2 G84864	probable membrane
13	575.5	20.7	524	2 A31318	Glucose transport
14	566.5	20.4	560	2 T51485	sugar transporter-
15	563.5	20.3	496	2 A41284	Glucose transport
16	560.5	20.2	493	2 A41751	Glucose transport
17	557.5	20.1	457	2 E70070	metabolite transpo
18	556.5	20.0	511	2 A84537	probable sugar tra
19	553	19.9	509	2 A32101	Glucose transport
20	551.5	19.8	511	2 H84536	probable sugar tra
21	550	19.8	509	2 A33801	muscle-fat glucose
22	549	19.8	523	2 S06920	Glucose transport
23	548.5	19.7	493	2 S38981	Glucose transport
24	542.5	19.5	496	2 A31986	Glucose transport
25	542.5	19.5	522	2 A31556	Glucose transport
26	540	19.4	482	2 B69803	metabolite transpo
27	538	19.4	580	2 D86426	hypothetical prote
28	536.5	19.3	508	2 G84564	probable sugar tra
29	531.5	19.1	508	2 B30310	Glucose transport

30 531 19.1 473 2 G69789 sugar transporter
31 525 18.9 493 2 A85433 sugar transporter
32 523 18.8 464 2 AC0877 galactose-proton s
33 522 18.8 464 2 F65079 galactose-proton s
34 522 18.8 549 2 T14606 probable sugar tra
35 517.5 18.6 492 2 I45902 glucose transporte
36 517 18.6 461 2 D70073 metabolite transpo
37 515.5 18.5 464 2 C91106 galactose-proton s
38 515.5 18.5 464 2 F85951 galactose-proton s
39 515.5 18.5 513 2 T01506 probable hexose tr
40 512.5 18.4 492 2 A27217 glucose transport
41 512.5 18.4 492 2 A25949 glucose transport
42 511.5 18.4 492 2 S03705 glucose transport
43 511.5 18.4 523 2 S25015 monosaccharide tra
44 510.5 18.4 492 2 A30797 glucose transport
45 508 18.3 469 2 H97064 probable sugar-pro

ALIGNMENTS

RESULT 1

T14545
probable sugar transporter protein - beet
C;Species: Beta vulgaris (beet)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T14545
R;Chou, T.J.; Bush, D.R.
Plant Physiol. 110, 511-520, 1996
A;Title: Molecular cloning, immunochemical localization to the vacuole, and expression
A;Reference number: Z18131; MUID:96351183; PMID:8742332
A;Accession: T14545
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-490 <CHI>
A;Cross-references: UNIPROT:Q39416; EMBL:U43629; NID:G1209755; PIDN:AAB53155.1; PID:G12
A;Experimental source: tonoplast
C;Superfamily: Glucose transport protein
C;Keywords: transmembrane protein

Query Match 63.2%; Score 1757.5; DB 2; Length 490;
Best Local Similarity 66.2%; Pred. No. 9.5e-118;
Matches 337; Conservative 76; Mismatches 73; Indels 23; Gaps 3;

QY 39 MGGGSRGGAGAGEESGDHGVLRRLPLNTGSRWMSRQSSFAFGTSSMAVLRESHVS 98
Db 1 MSDSFAGLGGGG-----GDLKPFHTGSRWMSRQSSL---MGSSQVIRESSIS 49
QY 99 AFLCTLIVALGPIQFGFTSGFSPTQDAMVRDLNISFSAFGLSNVGMVGAIASQ 158
Db 50 VLACVLIVALGPIQFGFTAGYSPTQSAITNELGLSVAEYSWFGSLSNVGMVGAIASQ 109
QY 159 MAEYIGRKGSIMTAIPNIIGWLAIISFAKDAFLYMGELLEGVGIISYTVPVYIAETIS 218
Db 110 ISEYIGRKGSIMTAIPNIIGWLAIISFAKDSFLYMGELLEGVGIISYTVPVYIAETIS 169
QY 219 PQNMRGALGSVNQLSVTFGIFLAYLIGMPVRLRLAVIGALPCTMLIPGLFFIPESPRWL 278
Db 170 PQNMRGALGSVNQLSVTFGIFLAYLIGMPVRLRLAVIGALPCTMLIPGLFFIPESPRWL 229
QY 279 AKNULTDETSLOVLRGFETDITTEVNDIKAVASSSKRTTISFQELNKKYRTPLLIG 338
Db 230 AKNMMEEFEVSLQVLRGFDTDISLEVNIEKSVASSSKRTTIRFAELQRRYWLPLMTG 289
QY 339 IGLLVQLNGSINGVLFYASSIFKAAAGVTNSDLATCSLGAIQVLTATGVTTLLDRAGRRI 398
Db 290 NGLLIQLQSGINGVLFYSTIFKEAGVTSSNAATFGLGAVQVIATVTTTLVDKSGREL 349
QY 399 LIITSGMTCLLAVSVVFFLKDNISQDSNYYITLMISLVGIVSVFVTFPGMGAIAPW 458
Db 350 LLTVSSSGMTSLILVAVMSFFLKEMVSDSTWYSFSLSVGVGVAMVVTFLGLGAIAPW 409
QY 459 LWMSEILPVSIXSLGSGIATLANWLTSPAITTNTMLTWSVGGTFLSVWVSATIVFV 518

```

Db 410 IIMSEILPINIKLAGSIATLANFVAWIVTMTANIMLSNNSGTFIYMWVCAFTVAFV 469
QY 519 VLWVPETKGNRSRGTTFVLSISIORQLOW 547
Db 470 VIVWVPETKGR-----LEEQW 486

RESULT 2
E96782
hypothetical protein F22H5.6 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: E96782
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E96782
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-487 <STO>
A:Cross-references: UNIPROT:Q9FRL3; GB:A8005173; NID:g10052276; PIDN:AAG12689.1; GSPDB:G
C:Genetics:
C:Superfamily: glucose transport protein
A:Map position: 1
C:Superfamily: glucose transport protein

Query Match 62.4%; Score 1734.5; DB 2; Length 487;
Best Local Similarity 70.2%; Pred. No. 4-le-116;
Matches 334; Conservative 67; Mismatches 68; Indels 7; Gaps 2;

QY 52 BESSGDHGVLRRLNLTNGSWYRMSRQSSFPATQTSNAVLRESHVAFCTLIVAGPI 111
Db 8 BEARN-----LRRPFIHTGYSYRMSRQSSM---WGSQVIRDSISVLACVLIVAGPI 60

QY 112 QFGTSGSPSTQDAMVRDLNLS:SEFSAFGLSNVGMVCAIASGOMAYIGRKGLMI 171
Db 61 QFGTSGSPSTQAAIKDLGLTVSEYVFGSLSNVGMVCAIASGOMAYIGRKGLMI 120

QY 172 AAIPNIIGWLAIKSPAKDASFLYNGRLLEGFGVGIISYTPVPIAIBISPNQNRGALGSVQ 231
Db 121 AAIPNIIGWLAIKSPAKDASFLYNGRLLEGFGVGIISYTPVPIAIBISPNQNRGALGSVQ 180

QY 232 LSVTFGLFLAVLGNFPRLLAVIGALPCTMTLPGFFFPESPRWAKNLTEDCETSL 291
Db 181 LSVTIGIMLAVLLGLVFPWRILAVIGLIPCTLIPGFFFPESPRWAKNLTEDCETSL 240

QY 292 QVLKRGFTDITTEVNDIKRAVASSKRTTISFQELNOKKYRTPLLLGILGLVLQNLGIN 351
Db 241 QVLKRGFTDITVEVNEIKRSVASSTKNTVRFVLDKRRYVFLMVGIGLLVLQQLGIN 300

QY 352 GVLFPYASSIFKAAVWNSDLATCSLGAIOVLATGVTTWLLDRAGRRLLIISGMLTCL 411
Db 301 GVLFPYSSIFESAGVTSNAATFGVAGIOVATAISTWLVDKAGRRLLIISVGMVITSL 360

QY 412 LAVSVVFLKDNISQDSNYIILWISLVGIVSVITFSGMGAIPMLMSEILPVSIS 471
Db 361 VIVAAFLKFEVPSDSMYSWLSILSVGVVAMVVFSLGMPPLINSEILPYNKIG 420

QY 472 LGGSIATLANWLTGFATMTNMLTWSVGTFIUSYMWVSAFTIVFVWVPETKG 527
Db 421 LAGSIATLANWFFSWLITMTANLLAWSGGTFTLYGLVCAFTVVFVTLWVPETKG 476

RESULT 3

```

```

T52132
probable sugar transporter protein ERD6 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C:Accession: T52132
R:Kiyosue, T.; Abe, H.; Yamaguchi-Shinozaki, K.; Shinozaki, K.
Biochim. Biophys. Acta 1370, 187-191, 1998
A:Title: ERD6, a cDNA clone for an early dehydration-induced gene of Arabidopsis, encode
A:Reference number: Z5973; MUID:98213606; PMID:9545564
A:Accession: T52132
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-496 <KIY>
A:Cross-references: UNIPROT:O65799; EMBL:D89051; PIDN:BAAZ5989.1
C:Genetics:
C:Gene: ERD6
C:Superfamily: glucose transport protein
C:Keywords: transmembrane protein; transport protein

Query Match 33.1%; Score 920.5; DB 2; Length 496;
Best Local Similarity 41.4%; Pred. No. 4.7e-58;
Matches 186; Conservative 91; Mismatches 159; Indels 13; Gaps 3;

QY 98 SAFLCTLIVAGPIQFGFTSGFSPSTQDAMVRDLNLS:SEFSAFGLSNVGMVCAIASG 157
Db 57 SVFLSTFVAVSGSPCTGCGVGFSSGAQAGITKLSLVAEYSMEGSLITLGLIGAVFSG 116

QY 158 QMAEYIGRKGLMIAAIPNIIGWLAIKSPAKDASFLYNGRLLEGFGVGIISYTPVPIA 217
Db 117 KVADVLGRKRTTLCFCEFFCITGWLCAVLAQNAAMLDGRLGLLGIGVGFISYVPIA 176

QY 218 SPQNRGALGSVQNLISVTFGLVLLGWFIPWRLLAVIGALPCTMTLPGFFFPESPRW 277
Db 177 APKHVGRGSVFANQNLQNGCISLFFIIGNFIPWRLLTVGLVPCVHFVFCIFFPESPRW 236

QY 278 LAKNLTEDCETSLQVLRGFETDITTEVNDIKRAVASSKRTTISFQELNOKKYRTPLLL 337
Db 237 LAKLGRDREKSSQLRGRSDVDISREANTIRDTIDMTENGKTMSELFORRYAVPLII 296

QY 338 GIGLIVQLNSINGVLFYASSIFKAAVWNSDLATCSLGAIOVLATGVTTWLLDRAGR 397
Db 297 GVLGMLQQLCGSSGVTYASSLFNKGFP-SAGTSVIATIMVPKAMLATVLDKGR 355

QY 398 ILLIISTQSWTLCILAVSVVFPKDNISQDSNYIIL-----TMISLVGIVSVITFSGM 453
Db 356 TLLMASCASAMGSLALLSV-----SYGFSFGILPELPIPTCIGVLGHVFSFAMGM 407

QY 454 GAIPMLMSEILPVSISKLSGSIATLANWLTGFATMTNMLTWSVGTFIUSYMWVSAF 513
Db 408 GGLPWIIMAEIPPMVKVSAAGTIVTVNWLFGWITTYTFNFMLEWNSGMLIFSMVSAF 467

QY 514 TIVFVVLWVPETKGNRSRGTTFVLSISQ 542
Db 468 SIVFVFLVFPETKGRSLEIQAALLNSVQ 496

RESULT 4
D96589
hypothetical protein T22H22.15 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: D96589
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712

```

A;Accession: D96589
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-483 <STO>
A;Cross-references: UNIPROT:Q5ZVM0; GB:AE005173; NID:g3776591; PIDN:AAC64898.1; GSPDB:GN
A;Gene: T22H22.15
A;Map position: 1

Query Match 29.7%; Score 824; DB 2; Length 483;
Best Local Similarity 35.0%; Pred. No. 3.4e-51;
Matches 178; Conservative 92; Mismatches 139; Indels 100; Gaps 6;

QY 53 EGGSDHGVLRRLPLNTGSWY-----RMSR-----QSFAPGTSSMAVLR 94
DB 26 EDGVPSSLLRKYLSNNSTLSLLKYTKDKRKKRGEDEANLAPETS--LINKE 83
QY 95 SHVSA-----FLCTILVALGPIQGFSGFSPQDAMVRDLNLSIEFSAFGSLSNV 147
DB 84 NODSSATITTTLLTTFVAVSGFVFGSVPQSDUTKELNSVAEISLFGSILII 143
QY 148 GGMVGAISQMAEYIGRKGLMIAAIPNIIGWLAIISFAKDAFLYMGRLLEGFGVGIIS 207
DB 144 GAMIGAMSRIADMIGRATMGFSEMFILGLWLAIVLSKVAIWLVDGRFLVGYGMGVFS 203
QY 208 YTPVPIYIAETSPQMRGALSVNQLSVTFGIFLAYLLGMFIPWRLLAVALGALPCTMLIPG 267
DB 204 FVPEVYIAETPKLGRGFTVHQLLCLGVSVTYLLGSGFIGRILALIGMPCVVQMMG 263
QY 268 LFIPESPRWL-----AKNLTEDCETSLQVLRGFETDITTEVNDIKRAVASSKR 318
DB 264 LFVIPESPRWLKIIVRNKSNQAKVKEEFIALQRLGKESADISYESNEIKDYTRLTDL 323
QY 319 TTISFQELNOKKYRTPLLGLLVLQNLGSLGVNLFYASSIFKAAGVNTSLATSLGA 378
DB 324 SEGSIVDLPQPOVAKSLVWGLVQFGVGVNGIAFYASSIFESAGVSK----- 374
QY 379 IQVLATGVTTWLLDRAGRIILLIISTGMLCLLAVSVVFFLKNISQDSNSYIITMIS 438
DB 375 -----TGMIA 379
QY 439 LVGIVSVITFSGFQGAIPWLMSEILPVSIKSLGSIATLANWLTSFAITMTNMLTW 498
DB 380 MVVVQVYTSFSLGMMGIPWIMSEIPIIDIKSAGSLVTVWSVWGSWIIISFTFNLMMW 439
QY 499 SVGTFLSYVMVVSFTVFWLVWVPEK 527
DB 440 NPAGTFYVFATVCGATVIFVAKLYPETKG 468

RESULT 5
hypotheical protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: B86221
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, B.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: B86221
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-474 <STO>
A;Cross-references: UNIPROT:Q04036; GB:AE005172; NID:g2342693; PIDN:AAB70420.1; GSPDB:GN
C;Genetics:

A;Map position: 1

Query Match 29.5%; Score 820.5; DB 2; Length 474;
Best Local Similarity 36.6%; Pred. No. 6e-51;
Matches 173; Conservative 85; Mismatches 138; Indels 77; Gaps 5;

QY 98 SAFCLTILVALGPIQGFSGFSPQDAMVRDLNLSIEFSAFGSLSNVGMVGAIALSG 157
DB 51 SVFLSTFVAVSGSFCTGCGVGFSSGAQAGITKDLSLVAEYSMFGSILITLGLGAVFSG 110
QY 158 QMAEYIGRKGLMIAAIPNIIGWLAIISFAKDAFLYMGRLLEGFGVGIISVTPVYIAE 217
DB 111 KVADVLRK-----RNAMWLDGRLLLGIGVGIFSVIPVYIAE 150
QY 218 SPQMRGALSVNQLSVTFGIFLAYLLGMFIPWRLLAVALG----- 259
DB 151 APKVRGSGFVAVANQMCISLFFIIGNFIPWRLTLLVGMILLPRDCKMLKIFLNDRCNV 210
QY 260 -----PCTMLIPGLFFIPESPRMLAKMNLTEDETSQVLRGFETDITTEVNDIKRAVA 313
DB 211 IRTGLVPCVFHVCLFFIPESPRMLAKGRDKECSSLQRLRGSDVDISREANTIRDTID 270
QY 314 SSSKRTTISFQELNOKKYRTPLLGLLVLQNLGSLGVNLFYASSIFKAAGVNTSLAT 373
DB 271 MTENGGETKMSLFFORVAYPLIIGVGLMFLQQLCGSSGVTYASSLFFNKGGFP--SAIGT 329
QY 374 CSLGAIQVLATGVTTWLLDRAGRIILLIISTGMLCLLAVSVVFFLKNISQDSNSYI 433
DB 330 SVIATIMVPMKAMLATVLDKGRRLM-----SFGI 361
QY 434 L-----TMISLVGIVSVITFSGFQGAIPWLMSEILPVSIKSLGSIATLANWLTSFAIT 489
DB 362 LPFLPTIFTCIGVLGHIVSFGAMGGLPWIIIMAEIFPMNVKVSAGTLVTVTNLWFGWIT 421
QY 490 MTNMLTWSVGTFLSYVMVVSFTVFWLVWVPEKXNSRGDTIFVSLSIQ 542
DB 422 YTFNFMLEWNASGMLIFSMVSASSIVFIYFLVPETKGRSLEEIQALLNNSVQ 474

RESULT 6
A86221
hypotheical protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: A86221
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, B.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: A86221
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-454 <STO>
A;Cross-references: UNIPROT:Q04037; GB:AE005172; NID:g2342687; PIDN:AAB70413.1; GSPDB:GN
C;Genetics:

A;Map position: 1

Query Match 29.1%; Score 808; DB 2; Length 454;
Best Local Similarity 37.7%; Pred. No. 4.4e-50;
Matches 163; Conservative 95; Mismatches 152; Indels 22; Gaps 5;

QY 92 LREHVSAA--FLCTILVALGPIQGFSGFSPQDAMVRDLNLSIEFSAFGSLSNVGG 149
DB 21 INECRITAVLVLFSTFFVSVGSGFCGCAAGYSVAQGIINDLGLSLVAQSMFGSINTFG 80
QY 150 MVGAISQMAEYIGRKGLMIAAIPNIIGWLAIISFAKDAFLYMGRLLEGFGVGIISVT 209

```

Db      81 MGAIFSGKADLMGRKGTMMFAQIFCFGWVAVALAKDSNWLIDGRSLSTGFAVGLLSYV 140
QY      210 VPVVIAEISPNMREGALGSVNQLSVTFGIFLAYLLGMFIPWRLLAVGALPCTMLIPGLF 269
Db      141 IPVIAEITPKHVGAFVFNQMLQSCGLSFFYVGNFVHRNIALGLIPCALQVVTLF 200
QY      270 FIPSPRLAKMNLTEDETSQVLRGFETDITTEVNDIKRAVASSSKRTTISFQELNOK 329
Db      201 FIPSPRLGLKGWHEKCRASLQSLRGDDADISEANTIKETMLFDEGPKSRVMDLQF 260
QY      330 KYRPELLIGLILVQLNLSINGVLFYASSIFKRAAGVNTNSDLATCSLGAIOVLATGVTW 389
Db      261 RYAPSVVIGVGLMLLQQLSGSSGLMYVGVDFKGGFPSS-IGSMILAVIMIPKALLGLI 319
QY      390 LLDRAGRRIILII-----STGWTLLCLLAVSVVFFLKNISQDSNSYVIL-----TWIS 438
Db      320 LVEKGRRLLLNDLYLQASTGCMCFSSLSPFCFR-----SYGMDELTPIFT 371
QY      439 LVGIVSVTFSPFGKGAIPWMMSEILPVSIKSLGSIATLANWLTSFALTMTNMLTW 498
Db      372 CIGVVGFISSFAVGMGLPIIXSEIFPMNVKVSAGTLVTLANWSFGWIVAFAYNEWLEW 431
QY      499 SVGGTFLSYMVV 510
Db      432 NASINTYYVII 443

RESULT 7
G85059
probable sugar transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: G85059
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: G85059
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-461 <STO>
A:Cross-references: UNIPROT:Q9M100; GB:NC_001268; NID:g7267233; PIDN:CAB80840.1; GSPDB:G
C:Genetics:
A:Gene: AT4G04750
A:Map position: 4
C:Superfamily: glucose transport protein

Query Match 26.9%; Score 748.5; DB 2; Length 461;
Best Local Similarity 35.5%; Pred. No. 7.8e-46;
Matches 168; Conservative 91; Mismatches 171; Indels 43; Gaps 6;

QY      81 SFAPGTSSMAVLRESHUSA---PLCTILVAGLPIQFGTSGFSSPTQDAWRDLNLSISE 137
Db      15 SASFNKSSLSLSEISNASTRPFLAFTVGCGLSPGCT----- 53
QY      138 FSAPGSLNVGGMVGAIASQMAEYIGRKGLMIAAIPNIIGWLAISSPAKDAFLYMGRL 197
Db      54 FSFFGSLTVGLILGALICGLADLVGRVVTIWTNVLVGLWLAFAFAKVDLLDLGL 113
QY      198 LEGFGVGIISYTPVPIAIEISPNMREGALGSVNQLSVTFGIFLAYLLGMFIPWRLLAVTG 257
Db      114 LOGISVGISSYLPFIYISELAPNLKGAASSLMQLEFVGVGLSAFYALGTAVAWRSIALIG 173
QY      258 ALPCTMLIPGLFPIESPRWLAKMNLTEDETSQVLRGFETDITTEVNDIKRAVASSSK 317
Db      174 SIFSLVVLPLFLFPIESPRWLAKVGEKEVGVGLLSLRKAQSDVSEAAILEYTHVBO 233
QY      318 RTTIS--FQELNOKKYRTPLLLGIGLLVQLNLSINGVLFYASSIFKRAAGVNTNSDLATCS 375
Db      234 QDIDSRGFFKLFQKVALPLTIGWLISMPQLGGLNGYTFYDTIFTSTGV--SSDIGFL 292
QY      376 LGAIQVLATGVTWLLDRAGRRIILIIISGWTLLCLLAVSVVFFLKNISQDSNSYVIL 435

```

```

Db      293 TSIVQMTGGVLGVLLVDISGRSLLIFSOAGMFLGCLATAISFFLQKNCWETGT---P 348
QY      436 MISLVGIVSVTFSPFGKGAIPWMMSEILPVSIKSLGSIATLANWLTSFALTMTNML 495
Db      349 INALISVMVFGSYGLGMGPPIWIIASEIYPVDVKGAAGTVCNLTSSWLVTYSNFL 408
QY      496 LTWSVGGTFLSYMVVSATFIVFVVLWVETKQ-----XNSRGDTIF 536
Db      409 LQWSSTGTFMFATVMGLGFVFTAKLVETPKGSLEIQSAFTDSTSDSIF 461

RESULT 8
H85059
probable sugar transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: H85059
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: H85059
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-457 <STO>
A:Cross-references: UNIPROT:Q9M029; GB:NC_001268; NID:g7267234; PIDN:CAB80841.1; GSPDB:G
C:Genetics:
A:Gene: AT4G04760
A:Map position: 4
C:Superfamily: glucose transport protein

Query Match 26.2%; Score 729; DB 2; Length 457;
Best Local Similarity 35.6%; Pred. No. 1.9e-44;
Matches 161; Conservative 91; Mismatches 170; Indels 30; Gaps 5;

QY      80 SSFAPGTSSMAVLRESHVSAPLCTILIV-ALGPIQFGTSGFSSPTQDAWRDLNLSISEF 138
Db      11 SSTSLSLSLSEISNACTRPFVLAFIGSCGAFAGCI-----Y 49
QY      139 SAFGSLNVGGMVGAIASQMAEYIGRKGLMIAAIPNIIGWLAISSPAKDAFLYMGRL 198
Db      50 SLFGSILTVGLILGALICGLKLTDLVGRVKTITNILFVIGWFAFAFAKGVLLDLGL 109
QY      199 EGFGVGIISYTPVPIAIEISPNMREGALGSVNQLSVTFGIFLAYLLGMFIPWRLLAVIGA 258
Db      110 QGISIGISVYLGVPYITTEIAPRNLRGAASSPAQLFAGVGISVFYALGTIVAVRNMLTLC 169
QY      259 LPCTMLIPGLFPIESPRWLAKMNLTEDETSQVLRGFETDITTEVNDIKRAVASSSKR 318
Db      170 IPSLWVLPLFLFPIESPRWLAKVGEVEVAVLSSLRGEKSDVSEAAILEYTHVKQ 229
QY      319 TTI---SFQELNOKKYRTPLLLGIGLLVQLNLSINGVLFYASSIFKRAAGVNTNSDLATCS 375
Db      230 QDIDRGRFFKLFQKVALPSTIGVLLALPQLGGLNGYSFYDTISFTSTGV--SSDFGFI 288
QY      376 LGAIQVLATGVTWLLDRAGRRIILIIISGWTLLCLLAVSVVFFLKNISQDSNSYVIL 435
Db      289 TSVQVMEGGLTGLVLDVSGRRTLLVVSQAGMFLGCLTTAISFFLKENHCWETCT---P 344
QY      436 MISLVGIVSVTFSPFGKGAIPWMMSEILPVSIKSLGSIATLANWLTSFALTMTNML 495
Db      345 VLLFVSMVFGSYGSGMGPIMPWIIASEIYPVDVKGAAGTVCNLTSSWLVTYSNFL 404
QY      496 LTWSVGGTFLSYMVVSATFIVFVVLWVETKQ 527
Db      405 LQWSSTGTFMLFATVAGLGFVFTAKLVETPKG 436

RESULT 9
G86220
hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)

```

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: G86220
R/Theologias, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: G86220
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-378 <STO>
A/Cross-references: UNIPROT:O04040; GB:AB005172; NID:G2342688; PIDN:AB70414.1; GSPDB:GN
C/Genetics:
A/Map position: 1

Query Match 24.2%; Score 671.5; DB 2; Length 378;
Best Local Similarity 34.3%; Pred. No. 1.9e-40;
Matches 148; Conservative 66; Mismatches 125; Indels 93; Gaps 4;

QY 119 FSSPTQDAMVRDLNISSEFSAFGLSNVGMVGAISGQMAEYIGRK----- 166
DB 3 YSSPAQSKIMEGLSVADYSPTSVMTLGGMITAVFSGKISALVGRQVTSYYCWTGSG 62
QY 167 -----GSLMTAAIPNIIGMLAISFAKDSFLYMGELLEGFGVGIISYTVPVVIA 215
DB 63 LSEFEFFWYDPTWMSIDVCCIFGWLAVAFADHIMLTGRLFGFGVGLISYVVPVIA 122
QY 216 EISPNMKGALSGVNLQSVTFGIFLAYLLGMFIPWRLAVIAGLPCTMTLPGLFPFESP 275
DB 123 EITPTFRGGFGYSLQLLQCLGSLMFTGNFFHRTLLAISAFQVLCIFPESP 182
QY 276 RWLAKNLTEDCETSLQVLRGPTDITTEVNDIKRAVASSKRTTISFQELNOKKYRTP 335
DB 183 RWLWYGGDQHELVSLKLGNSDILKEAEIR----- 216
QY 336 LIGIGLLVLQNSGVLVYASSIFPKAAGTVNSDLATCSLGAIVLATGVTWLLDRAG 395
DB 217 -TGLGMLLQPCGSAASIAVAARIFDKAGFP-SDIGTITLAVILIPQSIWMLTVDRWG 274
QY 396 RRIILITSGMTCLLAVSVVFLKDNISQDSNYILTMISLVGVSVFVITPSFGMGA 455
DB 275 RRPLL-----MGVVSFGIGLGG 292
QY 456 IPWLMWSEILPVSIKSLGSIATLANWLTSPAITWNTMLMTWVGTFPLSYMVVSAFTI 515
DB 293 LPWIMSEIFPVNVKRTAGSLVMTSNWPFNMIYSENFMIQWSASGTYTIFSGVSLVTI 352
QY 516 VFVVLWVPETKG 527
DB 353 VFIWTLVPEKG 364

RESULT 10
F86220
hypothetical protein [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: F86220
R/Theologias, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: F86220
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-490 <STO>
A/Cross-references: UNIPROT:O04041; GB:AB005172; NID:G2342689; PIDN:AB70415.1; GSPDB:GN
C/Genetics:
A/Map position: 1

Query Match 22.5%; Score 626; DB 2; Length 490;
Best Local Similarity 30.6%; Pred. No. 4.6e-37;
Matches 165; Conservative 84; Mismatches 162; Indels 128; Gaps 12;

QY 53 ESGSDHGDVLRPLANTGSKYMRSSROSPAPGTSMAVLRESHVSAFLCT----- 103
DB 2 ESGS-----MKTPLVNNOEARSSSITCGLLSTSVAY-TGSPVYGACSSWITNMPQ 55
QY 104 --LIVALGPIQGFPTS-----GFSSTPTQDAMVRDLNISSEFSAFGLSN 146
DB 56 IGIFIFLIKICFEFVDVNLWPNFVHYHNLKMSYSPAQSKIMEELGLSVADYSPTSVMT 115
QY 147 VGMVGAISGQMAEYIGRKSLMTAAIPNIIGMLAISFA-----KDSAS 190
DB 116 LGMITRAFSKIAAVIGRRQTMIAVFCIFGWLAVAFAYFLIKLTFHLLWFIKDKM 175
QY 191 PLYMORLLEGFGVGIISYTVPIYIAISPNMKGALSGVNLQSVTFGIFLAYLLGMFIPW 250
DB 176 LLNIGRGFLGFGVGLISYVWPVYIAETPKAFRGFSFNQLQSGFISLMFTTGNFFHW 235
QY 251 RLLAVIGALPCTMTLPGLFPFESPRLV-----AKNLTEDCETSLQVLRGPTDITTE 304
DB 236 RFLAUSAIPCGIQMCLFPFESPRLVHTLKAMVGRELEVLTKRNGENDILEE 295
QY 305 VNDIKRAVASSKRTTISFQELNOKKYRTPLLIGLGLVLQNSGVLVYASSIFKAA 364
DB 296 AAER-----IGLGLMLLQPCGSAASIAVAARIFDPA 328
QY 365 GVTNSDLATCSLGAIVLATGVTWLLDRAGRLL-----IISTSGMTCLLAVSVVFLK 421
DB 329 GVA---IYCHVCRPMTPTTDFELFVRSVHMFENWPLLLSG-----LVILTFP-- 379
QY 422 DNISQDSNYILTMISLVGVSVFVITPSFGMGAIPWLMWSEILPVSIKSLGSIATLAN 481
DB 380 -----GVVLSFGIGLGLPWIWVSEVPPVNVKRTAGSLVTVSN 417
QY 482 WLTSPTAITWNTMLMTWSV-----GGTFLSYMVVSAFTIVFVVLWVPETKG 527
DB 418 WFFSWIILFNFNMQWSAFYYTNTKMLYDFRTYIFAGVLSMSPFVWTLVPEKG 476

RESULT 11
D84922
probable sugar transporter [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C/Accession: D84922
R/Rabin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayan, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: D84922
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-348 <STO>
A/Cross-references: UNIPROT:Q9ZU87; GB:AB002093; NID:G4249409; PIDN:AD13706.1; GSPDB:GN
C/Genetics:
A/Map position: 2

Query Match 22.3%; Score 621; DB 2; Length 348;

Db 389 L-----LNKFGMSVSVMAIAFLVSPFEIIPGPIPMVMAEFSQGPRAALIAAF 441
 QY 480 ANMLTSPAITMTNMLTWSVGTFLSYM---VVSFTVVFVVLWVPETKG 527
 Db 442 SNTCTNFIVACQYIADFC--GPYVFFLPAGVLLAFTL-FTFFKVPETKG 489

RESULT 14
 T51485
 sugar transporter-like protein - Arabidopsis thaliana
 N:Alternate names: protein T21H19 70
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
 C:Accession: T51485
 R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
 submitted to the Protein Sequence Database, August 2000
 A:Reference number: Z25394
 A:Accession: T51485
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-560 <SAT>
 A:Cross-references: UNIPROT:Q9LFL3; EMBL:AL391148
 A:Experimental source: cultivar Columbia; BAC clone T21H19
 C:Genetics:
 A:Map position: 5
 A:Introns: 86/1; 124/2; 161/1; 205/2; 238/3; 276/2; 307/3; 359/1; 408/1; 432/3; 462/2; 5
 A:Note: T21H19_70
 C:Superfamily: Glucose transport protein

Query Match 20.4%; Score 566.5; DB 2; Length 560;
 Best Local Similarity 29.4%; Pred. No. 9.4e-33;
 Matches 133; Conservative 95; Mismatches 179; Indels 45; Gaps 7;

QY 105 IVALPTIQFGTSFSSPTQDQVMDNLS-----ISPSAF 141
 Db 111 VACLGAILFGYHGVNAGALEYLAQDGIENVTLOGKYMHIHFTPPVNGVIVSLLA- 169
 QY 142 GLSNVGVMGAIASGQMAEYIGRKGSLMTAAIPNIICWLAISPAKDASFLYMGRLIEGF 201
 Db 170 -----GATVGSFTGALADKFGTRTFQDPAIPLAIGAFLCATAQSVQTMIVGRLLAGI 223
 QY 202 GVGIIISVTPVYIAEISPNQNRGALGVNQLSVTFGIFLAYLGMFIP-----WRLAVI 256
 Db 224 GIGISSAIVPLYIIEISPTIRGALGVNQLFICIGILAIALGLPLAANPLWMTWFGV 283
 QY 257 GALTCTMLIPGLFTIPSPRLAKMNLTEDETSLOVLRGFETDITTEVNDIKRAVASS 316
 Db 284 AVIPSVLLAIGMASPSPSPRLVQGGKVSBEAKIKLYKGR-VVELVRDLS-ASQGS 341
 QY 317 KRTTISFOELNQRKYRTPLLLGIGLLVQLNSINGVLFYASSIFKAAGVTNSDLATCSL 376
 Db 342 SEPEAGWFDLFSRRYKWKVSVGAALFLFQQLAGINAVVYSTSVFRSAGIQSDVAASALV 401
 QY 377 GAIQVLATGVTTWLLDRAGRRIILLIISTGNTLCLLAVSVVFLKDNISQDSNSYYILTM 436
 Db 402 GASNVFGTAVASSLMDKNGKSLSLTSPGGWALSLLLSLSTFWKALAAYSGT----- 454
 QY 437 ISLVGIVSFVITFGMGAIPWLMMSSEILPVSISKLSGSIATLANMLTSFAITM-TTNLM 495
 Db 455 LAVGVTVLYVLSFSLGAGVPALLPEIFASRIKAKAVALSLGMHWISNFVIGLYFLSVV 514
 QY 496 LTWSVGGTFLSYMVVSFTVVFVVLWVPETKG 527
 Db 515 TKFGISSVYLGFGVCLAVLYIAGNVVETKG 546

RESULT 15
 A41264
 glucose transport protein 3 - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 04-Sep-1998
 C:Accession: A41264
 R:White, M.K.; Rall, T.B.; Weber, M.J.

Mol. Cell. Biol. 11, 4448-4454, 1991
 A:Title: Differential regulation of glucose transporter isoforms by the src oncogene in
 A:Reference number: A41264; MUID:91342646; PMID:1875932
 A:Accession: A41264
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-496 <WHI>
 A:Cross-references: GB:M37785
 A:Superfamily: Glucose transport protein
 C:Keywords: transmembrane protein

Query Match 20.3%; Score 563.5; DB 2; Length 496;
 Best Local Similarity 29.6%; Pred. No. 1.3e-32;
 Matches 140; Conservative 95; Mismatches 199; Indels 39; Gaps 10;

QY 89 MAVLRSHVSAPLCTLIIVAGLPQFGTSFSSPTQDQVMDNLSISE-----F 138
 Db 1 MADKKKITASLIYAVSVAAGISLQGYNTGVINPEKIIQAFYNTLSQRGETISPLL 60
 QY 139 SAFGSL-----NVGVMGAIASGQMAEYIGRKGSLMTAAIPNIIG--WLAI-FAKDASF 191
 Db 61 TSLWLSVAIFSVMGWSVSLFVNRFRGRNSMLLVNVLAFAGCALMALSKIAKAVEM 120
 QY 192 LYNGRLLEGFGVGIISYVPIVIAISPNQNRGALGVNQLSVTFGIFLAYLGM----- 246
 Db 121 LIIGRFIIICLFCGLCTGFPVMYISEVSPSLRGAFGLNQLGIVVGLIIVAQIFGLEGIMG 180
 QY 247 -FIPWRLIIVAGLPCVTLIPGLFTIPSPRLAKMNLTEDETSLOVLRGFETDITTE 304
 Db 181 TEALWPLLGLFTIVPAVLCVALLFCPSPRELLINKKEEKAQIVLQKLRGTQ-DVSQD 239
 QY 305 VNDIKRAVASSSKRTTISFQEL-NOKKYRTPLLLGIGLLVQLNSINGVLFYASSIFKA 363
 Db 240 ISEMKEESAKMSQEKATVLELFRSPNYRQPIIISITLQLSQOLSGINAVFYSTGIFER 299
 QY 364 AGVTNSDLATCSLGAIOVLATGVTTWLLDRAGRRILLIISTGNTLCLLAVSVVFLKDN 423
 Db 300 AGITQPVYATIGAGVNTVTVVSLFLVERAGRRIHLVGLGMAVCRAVMTIALAKEK 359
 QY 424 ISQDSNSYYILTMISLVGIVSFVITFGMGAIPWLMMSSEILPVSISKLSGSIATLANWL 483
 Db 360 -----WIRYISIVATFGFVALFEIGPGPIPWFIVAELFSGQPRPAAMAVAGCSNWT 410
 QY 484 TSPAITMTNMLTWSVGGTFLSYMVVSFTVVFVVLWVPETKGN-----SRG 532
 Db 411 SNFLVGMFLPYAEKLCGPYVFLIFLVLFFIFIFTYFKVPETKGRFTEDISRG 463

Search completed: October 13, 2004, 11:39:11
 Job time : 16.714 secs

is Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 13, 2004, 11:40:19 ; Search time 4488.99 Seconds
(without alignments)
4489.025 Million cell updates/sec

Title: US-10-051-909-36

Perfect score: 2779

Sequence: 1 PSSSSSRPAGKKKKKNOG.....TIFVLSIORQLQWLPECLS 553

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool.h/US10051909/runat_13102004_123336_19917/app_query.fasta_1.1678
-DB=EST -QFMT=fastap -SUFFIX=1st -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=500 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptio -NORM=ext -HEAPSIZES=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051909 @CGN 1 1 6263 @runat_13102004_123336_19917 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB	ID	Description
1	1733.5	62.4	1837	3	CNSOACUK	EX815012	Arabidops
2	1219	43.9	854	7	CN125032	CN125032	RHOHL_8A
3	1102	39.7	868	6	CB675064	OSJNEB101	
4	1065	38.3	682	6	CA191028	SCCORT2C0	
5	1021	36.7	635	6	CA143053	SCQSRV203	
6	1004.5	36.1	764	6	CB648359	OSJNEB11L	
7	984.5	35.4	1155	7	CK211005	FGAS02283	
8	984	35.4	674	1	AV939950	AV939950	
9	976	35.1	714	6	CA262252	SCQLEB201	

10	975	35.1	681	1	AV913420	AV913420	
11	962.5	34.6	718	7	CR291514	CR291514	
12	962.5	34.6	762	1	AJ794429	AJ794429	
13	932	33.5	701	6	CD871121	AZ02_117H	
14	930.5	33.5	1675	3	CNSOABHD	Arabidops	
15	925	33.3	769	4	B1933646	EST535353	
16	919.5	33.1	1737	3	CNS092VC	Arabidops	
17	914	32.9	772	6	CB683315	OSJNEB11L	
18	907.5	32.7	1666	3	CNSOABHV	Arabidops	
19	907	32.6	666	5	BQ766951	EBR008_SQ	
20	904.5	32.5	627	6	CA180355	SCCST300	
21	904	32.5	627	6	CA201877	SCRPL102	
22	902	32.5	716	7	CF451024	EST687369	
23	901.5	32.4	748	6	CB347677	CAB2EG000	
24	896.5	32.3	1601	3	CNSOABG20	Arabidops	
25	896	32.2	685	5	BQ862009	OQC1412_Y	
26	893	32.1	566	4	BM325827	PIC1_53_C	
27	890	32.0	791	6	CB92803	EST645555	
28	887.5	31.9	746	6	CD483180	atc01_31m	
29	882	31.7	594	7	CD203945	HSL_2_E01	
30	882	31.7	772	7	CF835377	UCRCS03_0	
31	870.5	31.3	744	1	AJ796408	AJ796408	
32	869	31.3	641	6	CA246696	SCSGFL5C0	
33	864.5	31.1	927	7	CK153511	FGAS03213	
34	859	30.9	548	6	CA143695	SCRUT200	
35	850	30.6	627	4	BJ249193	BJ249193	
36	847	30.5	663	7	CN906216	010929ABD	
37	845	30.4	670	5	BQ855360	QGB23P02	
38	838	30.2	595	2	BE599181	P11_86_D0	
39	830.5	29.9	651	4	BG522368	20_64_Ste	
40	828	29.8	667	6	CD893209	G118_123C	
41	827	29.8	626	7	CN904481	01081ABD	
42	825	29.7	656	5	BQ913239	QMA620_Y	
43	822	29.6	1054	3	AY111571	Zea mays	
44	821	29.5	686	7	CO116723	GR_EB019	
45	820.5	29.5	714	6	CA262252	SCQLEB201	

ALIGNMENTS

RESULT 1
CNSOACUK 1837 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLTIS24ZB10 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION EX815012 GI:42472886
VERSION EX815012.1
KEYWORDS HTC; GSIT CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
REFERENCE 1 (bases 1 to 1837)
AUTHORS Castell, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
Temple, G., Caboche, M., Weissenbach, J. and Salancub, M.
TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1837)
AUTHORS Genoscope.
JOURNAL Direct Submission
TITLE Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
COMMENT - Web : www.genoscope.cns.fr
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,

Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

FEATURES

Location/Qualifiers
1. 1837
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTUS242B10"
/tissue_type="Adult vegetative tissue"
/plasmid="pCMVSPORT 6"
complement(1. 1837)
/gene="At1g19450"

ORIGIN

Alignment Scores:

Pred. No.: 1.6e-168 Length: 1837
Score: 1733.50 Matches: 338
Percent Similarity: 83.23% Conservative: 69
Best Local Similarity: 69.12% Mismatches: 74
Query Match: 62.38% Indels: 9
DB: 3 Gaps: 2

US-10-051-909-36 (1-553) x CNSOACUK (1-1837)

Qy 52 GluGluSerGlySerArgHisAspGlyValLeuArgArgProLeuLeuAnThrGlySer 71
Db 213 GAAGAGGGAAGAATGAT-----CTTCGGGACCGCTTCTACACAGGAGT 260
Qy 72 TrpTyrArgMetSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal 91
Db 261 TGGTATCGATGGGTCGAGACAACTAGTAGT-----TTGGAATCGTCTCAAGTT 311
Qy 92 LeuArgGluSerHisValSerAlaPheLeuCysThrLeuLeuValAlaLeuGlyProIle 111
Db 312 ATTGAGACAGCTCAATTCCTGCTTAGCTTGCTTTGATTTGCTCTCTGCTCTATT 371
Qy 112 GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu 131
Db 372 CAATTGGGATTCACCTTGTTGGTTATTCTCCAACTCAAGCTGCACTTAAGGATCTT 431
Qy 132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyMetVal 151
Db 432 GGTAACTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 491
Qy 152 GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgGlyGlySerLeuMetIle 171
Db 492 GGACCAATTCGAGTGGTTCAGATTGCTGATACCTTGACGGAAGGCTCTCTGATGATT 551
Qy 172 AlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLeuAspAlaSerPhe 191
Db 552 GCTGCAATTCCTCAATATTATTGGATGGCTTCGATATCATTTGCGAAAGATCTCTCTTT 611
Qy 192 LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro 211
Db 612 CTTTACATCGGAAGATTGTTAGAGGCTTCGAGTTCGGATTATCTCTATACGGTCCC 671
Qy 212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231
Db 672 GTATATATAGCTGATGCTGCTCCACAGACCATGAGAGGAGGCTTAGCTTCACTAAC 731
Qy 232 LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArg 251
Db 732 CTTTCTGTAACAATTGGGATTAATCTCGGCTATTATCTCGGCTCTCTTGTTCATGGAGA 791
Qy 252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPheIle 271

Db 792 ATCTTTCAGTTTGGGAGTATTGCCATGTACATTGTTGTATACCGGGTCTTTTTCATT 851
Qy 272 ProGluSerProArgTrpLeuAlaIleMetAsnLeuThrGluAspCysGluThrSerLeu 291
Db 852 CCTGAATCTCTCGGTGGTTGGCAAGATGGGTTTGACAGATGATTTTGGAAACTTCATTG 911
Qy 292 GlnValLeuArgGlyPheGluThrAspIleThrGluValAsnAsp-IleLysArgAl 311
Db 912 CAAGTCTCTTCGGGATTTGAGACTGATATTACCGTTGAGTTTATGAAATCAAGAGATC 971
Qy 311 aValAlaSerSerSerLysArgThrThrIleSerPheGlnGlnLeuAsnGlnLysLys 331
Db 972 TGTGGCATCATCTAGCAAAAGCTTCTGCGATTCTGGTTCGAGCTCAAGCGCAGAGATA 1031
Qy 331 rArgThrProLeuLeuLeuGlyIleGlyLeuValLeuGlnAsnLeuSerGlyIleAs 351
Db 1032 CTATTTCACACTGATGGTGGTATAGGCTGCTGCATCTTCAACAACTTGGAGGATCA 1091
Qy 351 nGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLe 371
Db 1092 TGGTGTCTTCTTATTTCAGTACAAATTTTGAATCTGCAGGGTGACATCAAGTAATGT 1151
Qy 371 uAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrpLeuLe 391
Db 1152 GCGGACATTTGGAGTTGGCGTTGTTTCAGGTAGTGGCGACTGGGATAGCGATGGTGGT 1211
Qy 391 uAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLe 411
Db 1212 GGATAAGCAGGCTCGGCTTCTACTCATCATCTCTCTATCGAATGACGATAGGCT 1271
Qy 411 uLeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTy 431
Db 1272 CGTGATTTAGCAGTTGCATTTTACCTTAAGCAATTTGTATCACTCATTCACATGTA 1331
Qy 431 rTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPh 451
Db 1332 CAACATTCAGCATGGTTTCCGTTAGTTGGAGTTGGCTATGTTATCTCTGCTCTCT 1391
Qy 451 eGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLysSe 471
Db 1392 AGGAATGGACCAATTCATCGCTGATTATGCTGAGATTCTGCAGTGAATATAAAGG 1451
Qy 471 rLeuGlyLysSerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetTh 491
Db 1452 TTTACCCGGAAGTATAGCAACTTTTAACTGGTTCGTGCTGTTAGTGACATGAC 1511
Qy 491 rThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValSe 511
Db 1512 TCCAATATGCTGTAGCATGGAGCAGTGGGGAACCTTCTCTACGCTTTGGTTTG 1571
Qy 511 rAlaPheThrIleValPheValValLeuTrpValProGluThrLysGly***AsnSerAr 531
Db 1572 TGGATTCATCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1630
Qy 531 gGlyAspThrIlePheValSerLeu 539
Db 1631 AAGAGATCAAGCTTTGTCAGATG 1655

RESULT 2

LOCUS CN125032 854 bp mRNA linear EST 01-APR-2004
DEFINITION RHOH1_8_A07_g1_A002 Acid- and alkaline-treated roots Sorghum
bicolor cDNA clone RHOH1_8_A07_A002 5', mRNA sequence.
ACCESSION CN125032
VERSION CN125032.1 GI:45949053
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 854)
REFERENCE Cordonnier-Pratt M.-M., Suzuki, Y., Sugano, S., Klein, R.R., Liang, C.,

Sun, F., Sullivan, R., Lim, S., Eastman, A. and Pratt, L.H.
 An EST database from Sorghum: acid- and alkaline-treated roots
 Unpublished (2003)
 Contact: Cordonnier-Pratt MM
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: mmpratt@uga.edu
 Library constructed by Dr. Yuraka Suzuki and Dr. Sumio Sugano in
 the Human Genome Center, University of Tokyo. Institute of Medical
 Science; plant material and RNA prepared at Texas A & M University;
 sequencing done in the Laboratory for Genomics and Bioinformatics,
 University of Georgia. Sequence ends have been trimmed to exclude
 vector and regions below Phred quality 16. Three-prime sequences
 are presented as their reverse complement and have been trimmed to
 exclude polyA.
 Seq primer: Sug5 (CTTCTGCTCTAAAGCTGCG) .
 Location/Qualifiers
 1. .854
 /organism="Sorghum bicolor"
 /mol_type="mRNA"
 /cultivar="Brx623"
 /db_xref="taxon:4558"
 /clone="RHOH1.8 A07 A002"
 /lab_host="DH10B-T1 phage-resistant E. coli"
 /clone_lib="Acid- and alkaline-treated roots"
 /note="Organ: Root; Vector: pME18S-FL3; Site_1: XhoI;
 Site_2: XhoI; The library was prepared from polyA+ RNA
 from 8-day-old roots harvested from Brx623 sorghum
 seedlings grown in hydroponic culture. HCl was added to a
 pH of 3.0 to some seedlings, KOH to a pH of 9.0 for
 others. Roots were harvested 3, 12 and 27 hr after
 addition and pooled for RNA preparation. Double-stranded
 cDNA was cloned unidirectionally into different DraIII
 sites of the pME18S-FL3 vector [5-prime DraIII site is
 CACTGTGTG, 3-prime DraIII site is CACCATGTG]. XhoI excises
 the cDNA insert."

ORIGIN

```

Alignment Scores:
Pred. No.:      1.32e-115      Length:      854
Score:          1219.00      Matches:      250
Percent Similarity: 93.43%      Conservative: 6
Best Local Similarity: 91.24%      Mismatches:  11
Query Match:      43.86%      Indels:      7
DB:              7              Gaps:        3

US-10-051-909-36 (1-553) x CN125032 (1-854)

QY      3 SerSerSerPheArgProAlaGlyLyGlyLysLysLysAsnGlnGlyLeuArg 22
      |||
Db      51 TCTTCTCTCTCTTTCAGCCACGCGGGC--AAGAAGAAGAAGAAAAATCAAAAGCTTGGG 107
      |||

QY      23 -ArgGluAlaValProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlyGly 42
      |||
Db      108 GCGAGAGCGGTGCCCGCGGCAAGCAGCGAA-----GCCATGGGTGGCGGCGG 158
      |||

QY      42 YSerAsnArgGlyAlaGlyAlaGlyGluGluSerGlySerAspHisAspGlyValLe 62
      |||
Db      159 CAGCAACACAGAGGC-----GGCGCCACGAGAGAGCGCGCAGCACGCGGGCGCT 212
      |||

QY      62 uArgArgProLeuLeuAsnThrGlySerTrpTyraGMetSerSerArgGlnSerSerPh 82
      |||
Db      213 GCGGAAGCCGCTGCTCCACACGCGGCACGTGGTACCGCTATGAGCTGCGGCAGCTCAGCGT 272
      |||

QY      82 eAlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCy 102
      |||
Db      273 GCGCCCCCGGGCGCTCTCCATGCGCGGTCTCGGGAGTCCACAGCTCTCGCCTTCTCTG 332
      |||

QY      102 eThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerPr 122
      |||

```

Db	333	CACACTCATCTCGCGGCTTGGCCCACTCAGTTTCGGGCTTACACAGCGGCTTCTCTCTCC	399
Qy	122	oThrGlnAspAlaMetValArgAspLeuAnLeuSerIleSerGluPheSerAlaPheG	142
Db	393	GACCCAGGACACCATGGTGGGACCTCAACCTCTCCATCTCCAGTTCTCGCGGTTGG	452
Qy	142	ySerLeuSerAnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGlu	162
Db	453	CTCGCTGTCCAACTGTCGGCGCATGGTGGGGCGATCGCCAGCGGCGATGGCCGAGCA	512
Qy	162	rileGlyArgLyGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAl	182
Db	513	CATTGGCGCTAAAGGTTGTTGATGATTGCTCAATCCCGAATCATCGTTGGCTTGC	572
Qy	182	alleSerPheAlaIlyaspAlaSerPheLeuTyMetGlyArgLeuLeuGlyGlyPheG	202
Db	573	GATCTCCCTTTCGAAAGATGCTCATTTCTTTATATGGGACGATTCCTTGAAGGGTTGG	632
Qy	202	yValGlyIleIleSerTyThrValProValTyIleAlaGluIleSerProGluAsnMe	222
Db	633	TGTCGGCATCATATCCTATACGGTACCTGTATACATAGCAGATATCTCTCAGACAT	692
Qy	222	tArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTy	242
Db	693	GAGAGGAGCTCTTGTTCTGTGACACCAAGTTGTCGGTACCTTTGGCATATTTCTGGCCTA	752
Qy	242	rleuLeuGlyMetPheIleProTyArgLeuLeuAlaValIleGlyAlaLeuProCysTh	262
Db	753	TTTGTCGGCATGTTTGTCTCTGGAGACTTCTAGCTGTGATCGGAGCCTTGGCCCTGAC	812
Qy	262	rMetLeuIleProGlyLeuPhePheIleProGluSerPro	275
Db	813	AGTGTGATTCTCGACTATTTCTTCATTCAGAAATCTCCC	852
RESULT 3			
LOCUS	CB675064		
DEFINITION	OSUNEE10113.f OSUNEE Oryza sativa (japonica cultivar-group) cDNA		
ACCESSION	clone OSUNEE10113 5', mRNA sequence.		
VERSION	CB675064		
KEYWORDS	CB675064.1 GI:29678789		
SOURCE	EST.		
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE	Oryza sativa (japonica cultivar-group)		
AUTHORS	Oryza sativa (japonica cultivar-group)		
TITLE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
JOURNAL	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
COMMENT	Ehrhartoideae; Oryzaceae; Oryza.		
	1 (bases 1 to 868)		
	Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.		
	Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.		
	Large-scale identification of ESTs involved in the interaction		
	between rice and Magnaporthe grisea		
	Unpublished (2003)		
	Contact: Rod Wing		
	Arizona Genomics Institute		
	University of Arizona		
	Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ		
	85721-0088, USA		
	Tel: 520 626 3967		
	Fax: 520 621 9288		
	Email: http://genome.arizona.edu		
	PCR PRimers		
	FORWARD: gta aaa cga cgg cca gtc		
	BACKWARD: gga aac agc tat gac cat g		
	Plate: 10 row: I column: 13		
	Seq primer: gta aaa cga cgg cca gtc.		
FEATURES			
source	Location/Qualifiers		
	1..868		
	/organism="Oryza sativa (japonica cultivar-group)"		
	/mol_type="mRNA"		
	/cultivar="Nipponbare"		
	/db_xref="taxon:39947"		
	/clone="OSUNEE10113"		

```

/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNE"
/notes="vector: pBluescript II KS +; Site: 1: EcoRI; Site 2:
XhoI; 24 hrs after inoculation with Rice Blast (70-15)"

ORIGIN
Alignment Scores:
Pred. No.: 1.93e-103 Length: 868
Score: 1102.00 Matches: 208
Percent Similarity: 87.89% Conservative: 46
Best Local Similarity: 71.97% Mismatches: 35
Query Match: 39.65% Indels: 0
DB: 6 Gaps: 0

US-10-051-909-36 (1-553) x CB675064 (1-868)

QY 206 IleserTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAla 225
DB 1 ATATCTTATGTGGTCCAGTTTATATAGCAAGATTCTCCACAAACATGAGGAGCT 60

QY 226 LeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGly 245
DB 61 CTTGGTTTCAGTCAATCAGCTTCTGTACTATTGGCATATTGCTTACTTTGTAGGC 120

QY 246 MetPheIleProTyrArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIle 265
DB 121 ATGTTGTTCCTCGAAGATTCCTCGTCTCGTATTTCATGTTCAATCTGATA 180

QY 266 ProGlyLeuPhePheIleProGluSerProArgTyrPheLeuAlaLysMetAsnLeuThrGlu 285
DB 181 CCTGGATTGTTCTTATCCCTGAATCACCAGGTGGTGGCAAAATGGGAAGATGAAG 240

QY 286 AspCysGlnThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGluVal 305
DB 241 GATTTTGAATCCTCGTGCAGATTCGAGGATTTGAAACAGACATCGCTGTAGAGTA 300

QY 306 AsnAspIleLysArgAlaValAlaSerSerLysArgThrThrIleSerPheGlnGlu 325
DB 301 AATGAATAAGACAGACAGTTCATCATCATCGGAGGAGGACGACAAATACGATTTCGAGAT 360

QY 326 LeuAsnGlnLysLysTyrArgThrProLeuLeuGlyIleGlyLeuValLeuGln 345
DB 361 ATCAAGCAGACAGATATAGTGTAACCTTAATGATAGGAATTGGTCTCTGTACTGCAG 420

QY 346 AsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGly 365
DB 421 CAGTTAAGTGTGTGATGGCATCTTATTTATGCTGCTAGTATCTTCAAGCGCGGT 480

QY 366 ValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGly 385
DB 481 CTTACGAATAGTAATCTAGCAACATTTGGTCTGGGGTGTGTCAGGTGGTGTACTGGA 540

QY 386 ValThrThrTyrLeuLeuAspArgAlaGlyArgGlyIleLeuLeuIleSerThrSer 405
DB 541 GTGCAACTTGGTTGACGACAAAGCTGGTCGACACTCTCTCTCATATCTCTACTACA 600

QY 406 GlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeuLysAspAsnIleSer 425
DB 601 GGGATGACCATTAATCTTGTGTGTGTCTGTGTCTATTTTGTGAGGAGCAACATAACT 660

QY 426 GlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPhe 445
DB 661 AATGGTCTCTATTATATCTGATAGTATCTGATGATGATGATGATGATGATGATGAT 720

QY 446 ValIleThrPheSerPheGlyMetGlyAlaIleProTyrLeuMetMetSerGluIleLeu 465
DB 721 GTGATTTCAATTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780

QY 466 ProValSerIleLysSerLeuGlyIleGlySerIleAlaThrLeuAlaAsnThrLeuThrSer 485
DB 781 CCTGTAATAATCAAGAGCCTTGTGGGAAGCGTTCGACCCCTTGCACAACTGGTGTGACGGCT 840

```

```

QY 486 PheAlaIleThrMetThrThrAsnLeu 494
DB 841 TGGCTCATTACATGACGACGAGCTTG 867

RESULT 4
LOCUS CA191028 682 bp mRNA linear EST 24-SEP-2003
DEFINITION SCCRT2C03A08.9 RT2 Saccharum officinarum cDNA clone SCCRT2C03A08
5' mRNA sequence.
ACCESSION CA191028
VERSION CA191028.1 GI:35135776
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
REFERENCE
1 (bases 1 to 682)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
CONTACT: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parnada@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bccccenter.fcav.unesp.br
Plate: C03 row: A column: 08
Seq primer: T7 Promoter Primer.
FEATURES
Location/Qualifiers
1..682
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCCRT2C03A08"
/lab_host="DH10B"
/clone_lib="RT2"
/notes="Organ: Root tips(0.3cm-long) from adult plants;
Vector: pSport1; Site: 1: SalI; Site 2: NotI; An
unidirectional cDNA library generated from [Root
tips(0.3cm-long) from adult plants]. cDNA was prepared
from polyA+ mRNA using SuperScript Plasmid System Kit
(invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 Kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucest.lad.ic.unicamp.br/public"

```

```

ORIGIN
Alignment Scores:
Pred. No.: 9.18e-100 Length: 682
Score: 1065.00 Matches: 218
Percent Similarity: 97.36% Conservative: 3
Best Local Similarity: 96.04% Mismatches: 6
Query Match: 38.32% Indels: 1
DB: 6 Gaps: 0

US-10-051-909-36 (1-553) x CA191028 (1-682)

QY 256 IleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerPro 275
DB 2 ATCGGAGCGCTTGGCTCGACAGCTGTGTGATTCCTGAGCTATTCTTCATTCAGATCTCCC 61

QY 276 ArgTyrLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArg 295
DB 62 AGATGGCTGGCAAGATGAATTTGATGGAAGATTTCGAGACGTCCTACAAAGTGTGAGG 121

```

```

296 GlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSer 315
122 GGGTTTGAGACTGC-ATCAGCAGAGAGTGAATGATATAAGAGGGCAGTGACATCATCA 180
316 SerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeu 335
181 AGTAAGAGGACTACAAATCAGTTTTCAAGAATTAACCAAAAGAAATACCGCCGCGCTA 240
336 LeuLeuGlyIleGlyLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPhe 355
241 CTTCTAGGATTGGCCTTACTTGACTGCAAAATCTAAGTGGATCAACGGTGACTGTTT 300
356 TyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSer 375
301 TATGCAAGTAGCATCTTCAAGCTGCAGGTGTTACAAACAGCGACTTGGCCACTGTTC 360
376 LeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrProLeuLeuAspArgAlaGly 395
361 CTTGGAGCTATCCAGGTCCTTCTACTGGAGTTACGACATGGTTGTAGACAGCTGGA 420
396 ArgArgIleLeuLeuIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSer 415
421 CGAGCATGCTTCTCATTATTTACCTTGGCATGACTCTATGCCCTTCTTGGCTTCT 480
416 ValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThr 435
481 GTTGTATTTTCTCAAGGATAAGATTTCACAGGATTCTAAGCTGTAACCTTAACT 540
436 MetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAla 455
541 ATGATCTCCCTGGTGTGCTATCGTGGCTTTGTCAATACCTTCTCGTTGGTATGGTGCC 600
456 IleProThrLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlySer 475
601 ATTCATGGCTCATGATGCTGAGATCCTTCCAGTTAGCATCAAGATCTCGCGGAAGC 660
476 IleAlaThrLeuAlaAsnTrp 482
661 ATCCGACACTGGCCAAATGG 681

RESULT 5
CAL143053
LOCUS
DEFINITION SCOSRT2032A08.g RT2 Saccharum officinarum cDNA clone SCOSRT2032A08
5', mRNA sequence.
ACCESSION CAL143053
VERSION CAL143053.1 GI:35038703
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum

REFERENCE
1 (bases 1 to 635)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenhariaia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 032 row: A column: 08
Seq primer: T7 Promoter Primer.
FEATURES
Location/Qualifiers
source
1..635

```

```

/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCOSRT2032A08"
/lab_host="DH10B"
/clone_lib="Rr2"
/note="Organ: Root tips(0.3cm-long) from adult plants;
Vector: pSPori1; Site 1: SalI; Site 2: NotI; An
unidirectional cDNA library generated from [Root
tips(0.3cm-long) from adult plants]. cDNA was prepared
from polyA+ mRNA using SuperScript Plasmid System Kit
(Invitrogen). The double-strand cDNAs were fractionated
in a separose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 Kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucest.lad.ic.unicamp.br/public"

ORIGIN
Alignment Scores:
Pred. No.: 3,05e-95 Length: 635
Score: 1021.00 Matches: 204
Percent Similarity: 98.10% Conservative: 3
Best Local Similarity: 96.68% Mismatches: 4
Query Match: 36.74% Indels: 0
DB: Gaps: 0
US-10-051-909-36 (1-553) x CAL143053 (1-635)

QY 256 IleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPheIleProGluSerPro 275
DB 2 ATCGAGCGCTTGGCTGCACAGTGTTGATTCCTGGACTATTCTTCATCCAGAACTCCC 61
QY 276 ArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArg 295
DB 62 AGATGGCTGGCAAGATGAATTTGATGGAAGATTGCGAGACGTCCTCAAGTGTGAGG 121
QY 296 GlyPheGluThrAspIleThrGluValAsnAspIleLysArgAlaValAlaSerSer 315
DB 122 GGGTTTGAGACTGACATCACGACAGAGTGAATGATATATAAGAGGGCAGTGACATCATCA 181
QY 316 SerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeu 335
DB 182 AGTAAGAGGACTACATCAGTTTCAAGAAATTAACCAAAAGAAATACCGCCGCGCTA 241
QY 336 LeuLeuGlyIleGlyLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPhe 355
DB 242 CTTCTAGGAATTGGCCTTACTTGCTGCAAAATCTAAGTGGAAATCAACGGTGACTGTTT 301
QY 356 TyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSer 375
DB 302 TATGCAAGTAGCATCTTCAAGCTGCAGGTGTTACAAACAGCGACTTGGCCACTGTTC 361
QY 376 LeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrProLeuLeuAspArgAlaGly 395
DB 362 CTTGGAGCTATCCAGGTCCTTCTACTGGAGTTACGACATGGTTGTAGACAGAGCTGA 421
QY 396 ArgArgIleLeuLeuIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSer 415
DB 422 CGAGCATGCTTCTCATTATTTACCTTGGCATGACTCTATGCCCTTCTTGGCTTCT 481
QY 416 ValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThr 435
DB 482 GTTGTATTTTCTCAAGGATAAGATTTCACAGGATTCCTAAGCTGTAACCTTAACT 541
QY 436 MetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAla 455
DB 542 ATGATCTCTTGGTGTGCTATCGTGGCTTTTGTTCATTAACCTTCTCGTTGGTATGGTGCC 601
QY 456 IleProThrLeuMetMetSerGluIleLeuPro 466
DB 602 ATTCATGGCTCATGATGCTCAGATCCTTCCA 634

```

```

RESULT 6
CB648359
LOCUS
DEFINITION
OSJNEB1110.f OSJNEB Oryza sativa (japonica cultivar-group) cDNA
764 bp mRNA linear EST 08-APR-2003
OSJNEB1110.5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 764)
Jantaasuriyavat C., Lu, G., Gowda, M., Hatfield, J., Zhou, B., Mazur, E.,
Kudrna, D., Dean, R., Soderlund, C., Wing, R. and Wang, G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished (2003)
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtg
BACKWARD: gga aac agc tat gac cat g
Plate: 11 row: L column: 10
Seq primer: gta aaa cga cgg cca gtg.
Location/Qualifiers
1..764
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEB1110"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEB"
/note="vector: plusescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 24 hrs after inoculation with Rice Blast (Che
86061)"
ORIGIN
Alignment Scores:
Pred. No.: 2,12e-93 Length: 764
Score: 1004.50 Matches: 202
Percent Similarity: 87.40% Conservative: 13
Best Local Similarity: 82.11% Mismatches: 24
Query Match: 36.15% Indels: 7
DB: Gaps: 2
US-10-051-909-36 (1-553) x CB648359 (1-764)
Qy 28 GlyArgProAlaSerGluLeuArgThrArgValMetGlyGlySerAsnArgGlyGly 47
Db 35 GGAGCCGAGGCGGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 94
Qy 48 AlaGlyAlaGlyGluSerGlySerGlySerAspHisAsp-----GlyValLeuArg 63
Db 95 GCGTCTGTCGCGGAGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 154
Qy 64 ArgProLeuLeuAsnThrGlySerTyrThrArgMetSerSerArgGlnSerSerPheAla 83
Db 155 AAGCGCTGTGTGAACACGCGGAGCTGTGTACAGGATGGGATCGCGGCTCGCGGCC 214
Qy 84 ProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCysThr 103
Db 215 -----TCTTCATGGCGGCCATCGGAGATCCACCGTCTCCGCTTCCTCTGCACG 265

```

```

Qy 104 LeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProThr 123
Db 266 CTCATCGTCGCGCTCGGCCCATCAATTCGATTCACCGCGCTTCTCTCACCACC 325
Qy 124 GlnPheAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySer 143
Db 326 CAGGACGCCCATCCCGGACCTCAAGCTCTCCATCTCCGAGTCTCTCGGCTTCGGTTCG 385
Qy 144 LeuSerAsnValGlyGlyMetValGlyValAlaIleAlaSerGlyGlnMetAlaGluTyrIle 163
Db 386 CTGTCCACAGCTCGGCCCATCGTTCGGAGCGATCGCCAGTGGCGGAGATGGGAGTACATT 445
Qy 164 GlyArgGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTyrLeuAlaIle 183
Db 446 GCGCGGAAAGGCTCGTTCATATTCGCGCGGTCTCTAACATCATTTGGTTCGCTTGGCCATC 505
Qy 184 SerPheAlaLeuAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyVal 203
Db 506 TCCITTCGAAAAGACGCGCTCAATTTTATACATGGAGCGCTTCTTGAAGGTTTGGTGT 565
Qy 204 GlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArg 223
Db 566 GGTGTATATCATATACGTCGCGGATATACATAGCAGAGATATCTCATCAGACACAGA 625
Qy 224 GlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeu 243
Db 626 GGAGCACTTGGCTCGTGAAACAGTTCGCTTACCATTTGGTTCGCTTGGCTTATTTG 685
Qy 244 LeuGlyMetPheIleProTyrArgLeuLeuAlaValIleGlyAlaLeuProCysThrMet 263
Db 686 CTAGGCATGTTTGTCTTGGAGGCTGCTTCAGTGTAGGAGTATAGGAGCATCCCATGATG 745
Qy 264 LeuIleProGlyLeuPhe 269
Db 746 TTAATACCTGCTCTATTC 763
RESULT 7
CB211005
LOCUS
DEFINITION
aestivum cDNA, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Triticum aestivum (bread wheat)
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 1155)
Allard, F., Crosby, W. L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D.,
Genswein, B., Graf, R., Gulick, P., Hrycan, L. D., Larocche, A.,
Links, M. G., McCarthy, E. B., Monroy, A., Muzak, I., Nilsson, D.,
Penniket, C., Roach, J. L. and Sarhan, F.
Functional Genomics of Abiotic Stress In Wheat and Canola Crops
Unpublished (2003)
Contact: Wm L Crosby
Bioinformatics
University of Saskatchewan, Department of Computer Science
1C101 Engineering Building, 57 Campus Drive, Saskatoon,
Saskatchewan, S7N 5A9, Canada
Tel: 306 966 1769
Fax: 306 966 2033
Email: fgas@cs.usask.ca
This sequence is the direct result of the Base calling software
Phred (default parameters). It is the raw base calls. To aid in the
identification of the high quality insert the software Lucy
(default parameters) has been run on this sequence. Lucy identified
the region [44,808].
Plate: L5B024 row: L column: 10.
FEATURES
Location/Qualifiers
1..1155
source

```

/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone_lib="Triticum aestivum FGAS: Library 5 GATE 7"
/notes="vector: pCMV.SPORT6; Crown and developmental stages of spike formation in wheat cultivar Norstar. 4 mRNA populations were combined before constructing the library. The first mRNA population is from 1cm crown sections after 30 days of cold acclimation. The second is from 1cm crown sections after 11 days of deacclimation (before deacclimation plants were fully vernalized for 49 days). The third is from different developmental stages of spike formation (5 to 50mm) that still have not emerged from the leaf (dissection required). The last is from different developmental stages of spike and seed formation after having emerged from the leaf (visible). First strand synthesis in this library was done in the presence of methylated dCTP thereby protecting from internal cleavage with NotI."

ORIGIN

Alignment Scores:
Pred. No.: 4.88e-91 Length: 1155
Score: 984.50 Matches: 195
Percent Similarity: 82.83% Conservative: 51
Best Local Similarity: 65.66% Mismatches: 51
Query Match: 35.43% Indels: 3
DB: 7 Gaps: 0

US-10-051-909-36 (1-553) x CK211005 (1-1155)

Qy 244 LeuGlyMetPheLeuProTrpArgLeuLeuAlaValleGlyLeuProCysThrMet 263
Db 54 GTAGCATGTTTGTCTCTGAGAAATCTCTCAGTCTCTGCAATTTTACCTTCTCAATC 113
Qy 264 LeuLeuProGlyLeuPheLeuProGlySerProArgTrpLeuAlaLysMetAsnLeu 283
Db 114 CTGATACCTGGTGTCTTCATCCCGAATCACAAGGTGCTGGCAAAATGGGAAG 173
Qy 284 ThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThr 303
Db 174 ATGAGGATTTGAATCTTCGTCAGGTCTCTGCGGATTTGAATGATATCTCAGCA 233
Qy 304 GluValAsnAspIleLysArgAlaValAlaSerSerLysArgThrThrIleSerPhe 323
Db 234 GAAGTAATGAATAAAGATATAGTTCCTCATCTAGGAGGAGCAACATACGATTT 293
Qy 324 GlnGluLeuAsnGlnLysLysTrpArgThrProLeuLeuLeuGlyIleGlyLeuVal 343
Db 294 GCAGAGATCAACACAGAGATATAGTTCCTCTTATGATAGAAATGGTCTCTTATA 353
Qy 344 LeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAla 363
Db 354 CTTGAGCAACTAAGTGTGTCAATGGCATCTCTTCTATGCTGCAAGTATCTTCAAGCT 413
Qy 364 AlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAla 383
Db 414 GCTGCTCTTAAAAATAGTAATCTAGCAACATGTGTTGGGGGCTGTTCAGTGTGCT 473
Qy 384 ThrGlyValThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIleSer 403
Db 474 ACTGGAATCACAACCTGTTGACTGACAAAGCTGGTGACGGCTACTTCTCATTTATCT 533
Qy 404 ThrSerGlyMetThrLeuCysLeuLeuAlaValSerValPhePheLeuLysAspAsn 423
Db 534 GCTGTAGGGATGACAGTCATCTCTGCTGTTCTGTTCTGTTCTGTTGTTGTTGTTGTT 593
Qy 424 IleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGlyIleVal 443
Db 594 ATAGGAGAGCTTCTCATTTACACTCTGATGATGATGCTTTTCTACCTGGCTGGCTGTG 653
Qy 444 SerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGlu 463

Db 654 GCATTTGTGCTTCCCTTTCCATGGAGGCCATCCCATGATCATTAATGTCTGAG 713
Qy 464 lleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeu 483
Db 714 ATCTTCTGTTTAACTCAAGATCTGCGGGAAGACCCACCTCGCAACTGGATG 773
Qy 484 ThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGlyThr 503
Db 774 ACCTCTGCTCATCAGCATGACCGGAGCTTGATGCTTAACTGGAGCAACGAGGAACC 833
Qy 504 PheLeuSerTyrMetValValSerAlaPheThrIleValPheValValTrpValPro 523
Db 834 TTGCTATA-TTCGCGCGGTTCATGGGACCCCTCTCTCTGCTGCTGCTGCTGCG 892
Qy 524 GluThrLysGly***AsnSerArgGlyAspThrIlePheValSerLeuSer 540
Db 893 GAGACCAAGGAAG-AACATTCGAGAAA-TCGGTTTTCGTTCCGCTGAAC 941

RESULT 8

AV939950

LOCUS

DEFINITION

AV939950

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. 674

/organism="Hordeum vulgare subsp. spontaneum"

/mol_type="mRNA"

/strain="H602"

/sub_species="spontaneum"

/db_xref="taxon:77009"

/clone="bah24m12"

/tissue_type="top three leaves"

/dev_stage="adult, heading stage"

/clone_lib="K. Sato unpublished cDNA library, strain H602 adult, heading stage top three leaves"

ORIGIN

Alignment Scores:

Pred. No.: 2.34e-91 Length: 674

Score: 984.00 Matches: 193

Percent Similarity: 91.93% Conservative: 12

Best Local Similarity: 86.55% Mismatches: 17

Query Match: 35.41% Indels: 1

DB: 1 Gaps: 0

US-10-051-909-36 (1-553) x AV939950 (1-674)

Qy 53 GluSerGlySerAspHisAspGlyValLeuArgArgPro-LeuLeuAsnThrGlySerTr 72

Db 3 GAGAGCGGAGGACCATCACACGCGGNAAGCAGCGCTGCTGGTCAACACGCGGAGCTG 62

Qy 72 pTyrArgMetSerSerArgGlnSerPheAlaProGlyThrSerSerMetAlaValle 92

```

Db      63  GTACGGATGGGTGGCGGCCAGTCCAGCCTCACGGCGGGCACCTCTCCATGGCCATCAT 122
Qy      92  uArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGl 112
:      :      :      :      :      :      :      :      :      :      :
Db     123  GCGGAGTCCACAGTCTCCGCTTCCTCTGTCACCATGATCGTCGGCTCGCGCCCATCCA 182
Qy     112  nPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeuAs 132
:      :      :      :      :      :      :      :      :      :      :
Db     183  GTTCGGTTTACCGCGCGCTCTCTCCGCCACCCAGGACGCCATCATCCGCGCACTCAA 242
Qy     132  nLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGl 152
:      :      :      :      :      :      :      :      :      :      :
Db     243  CCTCTCCATCTCCGAGTTCTCCGTTTCGGCTCGCTGCCAAGCTCGCGCCCATGGTCGG 302
Qy     152  yAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgGlySerLeuMetIleAl 172
:      :      :      :      :      :      :      :      :      :      :
Db     303  GGCCATGCCAGCGCCAGATGGCGGACACATTGGCGCGCAAGGGTCACTGATGATTGC 362
Qy     172  aAlaIleProAsnIleIleGlyTyrLeuAlaIleSerPheAlaIleAspAlaSerPheLe 192
:      :      :      :      :      :      :      :      :      :      :
Db     363  TGCATATCTTACATCATTCGGCTGGCGCACTCTCTTCGCAAGACACTTCTTTTCT 422
Qy     192  uTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValProVa 212
:      :      :      :      :      :      :      :      :      :      :
Db     423  GTATATGGAGCAATGCTCGAAGGATTGGTGTGGTGTGTCATATCTTACACGGTGCAGT 482
Qy     212  lTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLe 232
:      :      :      :      :      :      :      :      :      :      :
Db     483  ATACATAGCAGAGATTCTCTCTAGAACATGAGAGGGCGCTTAGGCTCTGTGAACAGTT 542
Qy     232  userValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProThrArgLe 252
:      :      :      :      :      :      :      :      :      :      :
Db     543  GTCGGTAACGATTGGTATCGTGTGGCTACATCTCGGCATGTTTCTTCTGGAGGAT 602
Qy     252  uLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIlePr 272
:      :      :      :      :      :      :      :      :      :      :
Db     603  GCITTCAGTGATAGGAATCTTGCAATGACAAATATTGATACCTGGTGTCTTCTTCATCC 662
Qy     272  oGluSer 274
:      :      :      :      :      :      :      :      :      :      :
Db     663  CGAATCT 669

RESULT 9
CA262285
LOCUS   SCEQB2019E02.g LB2 Saccharum officinarum cdna clone SCEQB2019E02
DEFINITION
5', mRNA sequence.
ACCESSION
CA262285
VERSION
CA262285.1 GI:35954083
KEYWORDS
EST.
SOURCE
Saccharum officinarum
ORGANISM
Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
REFERENCE
1 (bases 1 to 714)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCSEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenhariaia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bccccenter.fcav.unesp.br
Plate: 019 row: E column: 02
Seq primer: T7 Promoter Primer.
Location/Qualifiers

```

```

source
1..714
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCEQB2019E02"
/lab_hosts="DH10B"
/clone_lib="LB2"
/notes="Organ: Lateral buds from plants adult plants
growing in greenhouse; Vector: pSport1; Site 1: SalI;
Site 2: NotI; An unidirectional cDNA library generated
from [lateral buds from plants adult plants growing in
greenhouse]. cDNA was prepared from polyA+ mRNA using
SuperScript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-column and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucsest.lad.ic.unicamp.br/public"

ORIGIN
Alignment Scores:
Pred. No.: 1.74e-90 Length: 714
Score: 976.00 Matches: 188
Percent Similarity: 90.52% Conservative: 22
Best Local Similarity: 81.03% Mismatches: 22
Query Match: 35.12% Indels: 0
DB: 6 Gaps: 0

US-10-051-909-36 (1-553) x CA262285 (1-714)
Qy      92  LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIle 111
:      :      :      :      :      :      :      :      :      :      :
Db     15  CTCGCGAGCTCGCCATCTCGGCCACGTTCTGCAAGCTCATTTGGCTCGAGGTGCCATC 74
Qy     112  GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu 131
:      :      :      :      :      :      :      :      :      :      :
Db     75  CAGTTCTGCTTTTACCAGCGGGTACTCTCTCCCGCCAGCAGGACGCCATCATTTGTCGAC 134
Qy     132  AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal 151
:      :      :      :      :      :      :      :      :      :      :
Db     135  GGCTCTCTCCCTCTCGAGTTCTCGCTCTTCGGATCGTTATCAACAGTAGGGCATGGTA 194
Qy     152  GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgGlySerLeuMetIle 171
:      :      :      :      :      :      :      :      :      :      :
Db     195  GGTGCGCATCTCCAGTGGGCAACTTCCAGAGTATATCGTCTCGCAAGGGGTCTCTCATGATC 254
Qy     172  AlaAlaIleProAsnIleIleGlyTyrLeuAlaIleSerPheAlaIleAspAlaSerPhe 191
:      :      :      :      :      :      :      :      :      :      :
Db     255  GCTGCATTCCAACATCAATTTGGTGGCTCGCATATCATTTCCCAAGGATTCCTCTTTC 314
Qy     192  LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro 211
:      :      :      :      :      :      :      :      :      :      :
Db     315  TTGTTTATGGTTCGGCTGCTAGAGGATTTCGAGTCGGTGTAAATATCTTATACAGTACCG 374
Qy     212  ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231
:      :      :      :      :      :      :      :      :      :      :
Db     375  GTTTATATATGCAGAAATCGCTCTCAAGATCAGAGGGGAGTTCTTGGTCTGTCAATCAG 434
Qy     232  LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProThrArg 251
:      :      :      :      :      :      :      :      :      :      :
Db     435  CTCCTCCGTCACAATTTGGTATATTCCTTGGCTACCTGTTTGGCATGTTTGTCCCTGAGA 494
Qy     252  LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIle 271
:      :      :      :      :      :      :      :      :      :      :
Db     495  ATTCTTGTGTTCTAGGCGTTTACCTTGTCTTCAATCTCTGATTCCTGACATGTTCTTGT 554
Qy     272  ProGluSerProArgTyrLeuAlaIleMetAsnLeuThrGluAspCysGluThrSerLeu 291
:      :      :      :      :      :      :      :      :      :      :
Db     555  CTGATATCCCAAGTGGTGGCAAAATGGGAGAGATGGAGGATTTTGAATATTCACTG 614
Qy     292  GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAla 311
:      :      :      :      :      :      :      :      :      :      :
Db     615  CAAGTTCTTCGAGGATTTTCAGACGGCATCACACAGAGTAAATGAATAAAGAGATCA 674

```


QY 312 ValAlaSerSerSerLysArgThrThrIleSerPhe 323
 Db 675 GTAGCATCATCAAGGAGAGACACACATAATGTTT 710
 RESULT 10
 AV913420
 LOCUS
 DEFINITION
 AV913420 K. Sato unpublished cDNA library, cv. Haruna Nijo
 germination shoots Hordeum vulgare subsp. vulgare cDNA clone
 bags22a03 5', mRNA sequence.
 AV913420
 AV913420.1 GI:18209197
 VERSION
 EST.
 KEYWORDS
 SOURCE
 ORGANISM
 Hordeum vulgare subsp. vulgare
 Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooidae; Triticeae; Hordeum.
 1 (bases 1 to 681)
 REFERENCE
 Sato, K., Saitoh, D. and Takeda, K.
 Barley EST sequencing project in NIG and Okayama Univ
 Unpublished (2002)
 JOURNAL
 COMMENT
 Contact: Tadashi Shin-i
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp.
 FEATURES
 Location/Qualifiers
 1..681
 /organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="Haruna Nijo"
 /sub_species="vulgare"
 /db_xref="taxon:112509"
 /clone="bags22a03"
 /tissue_type="shoots"
 /dev_stage="germination"
 /clone_lib="K. Sato unpublished cDNA library, cv. Haruna
 Nijo germination shoots"
 ORIGIN
 Alignment Scores: 2.05e-90 Length: 681
 Pred. No.: 975.00 Matches: 191
 Score: 91.89% Conservative: 13
 Percent Similarity: 86.04% Mismatches: 16
 Best Local Similarity: 35.08% Indels: 2
 Query Match: 1 Gaps: 2
 DB: 1
 US-10-051-909-36 (1-553) x AV913420 (1-681)
 QY 53 GluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeu---AsnThrGlySer 71
 Db 18 GAGGGAGAGAGAGACCATGAC---ACGGGAAGCAGCGCTGCTGTCTCAACACCGGGAGC 74
 QY 72 TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal 91
 Db 75 TGGTACCAGGATGGGTGCGGCAGTCCAGCTCAGCGGGGACACTCTCTCATGGCCATC 134
 QY 92 LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIle 111
 Db 135 ATGGCGGAGTCCACGCTCTCGGCTTCTCTGCACCATGATGCTCGCGCTCGGCCCATC 194
 QY 112 GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu 131
 Db 195 CAGTTTCGGCTTCCACCGCGGGTCTTCCTCCCGCCAGGACGCATCATCCGGACCTC 254
 QY 132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal 151
 Db 255 AACCTCTCACTCCGAGTTCTCCGTGTTCCGCTGCTGTCACAGCTCGGGCCCATGTC 314

QY 152 GlyAlaIleAlaSerGlyGlnMetAlaClnTyrIleGlyArgLysGlySerLeuMetIle 171
 Db 315 GCGCCATCGCCAGCGGCAGATGCCGAGACATGGCCGCAAGAGGTCTACTGATGATT 374
 QY 172 AlaAlaIleProAsnIleGlyTyrLeuAlaIleSerPheAlaLysAspAlaSerPhe 191
 Db 375 GCTGCTATCTTACATCATCGGCTGGCTGGCCATCTCTTCGCAAAAGACACCTCTTTT 434
 QY 192 LeuTyrMetGlyArgLeuLeuGlyPheGlyValGlyIleIleSerTyrThrValPro 211
 Db 435 CTGTATATGGGACGATTGCTCGAAGGATTGTTGTGTGTGCATATCTACACGGTGGCA 494
 QY 212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231
 Db 495 GTATACATAGCAGAGATTTCCTCAGAACATGAGAGCGCTCTAGGCTGTGTGAACAC 554
 QY 232 LeuSerValThrPheGlyIlePheLeuAlaTyrIleLeuGlyMetPheIleProTyrArg 251
 Db 555 TTGTGCGTAACGATTGGTATCGTGTGGCTACATCTCGGCATGTTTGTTCCTTGGAGG 614
 QY 252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIle 271
 Db 615 ATGCTTGCACTGATAGGAATCTTCCATGACATATGATACCTGGCTCTGTCTTCATT 674
 QY 272 ProGlu 273
 Db 675 CCGGAA 680
 RESULT 11
 CR291514
 LOCUS
 DEFINITION
 Y745ellp5, mRNA sequence.
 CR291514
 CR291514.1 GI:44678080
 VERSION
 EST.
 KEYWORDS
 SOURCE
 ORGANISM
 Oryza sativa
 Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 1 (bases 1 to 718)
 REFERENCE
 Han, B., Feng, Q., Huang, Y. C., Ying, K., Li, Y., Guan, J. P., Zhu, J. J.,
 Zhao, Q., Hu, X., Liu, Y. L., Mu, J., Yu, Z., Chen, L., Fan, D. L.,
 Weng, Q. J., Zhang, L., Lu, Y. Q., Yu, S. L., Liu, X. H., Lu, T. T.,
 Zhang, Y. J., Lu, Y., Li, C., Li, T., Zhang, Y., Hu, H., Jia, P. X.,
 Zhang, L., Lan, L. F., Chen, W., Wu, S. A. and Xue, Y. B.
 Rice cDNA EST clone
 Unpublished (2003)
 JOURNAL
 COMMENT
 Contact: Han Bin
 National Center for Gene Research
 Chinese Academy of Sciences
 500# Cao Bao Road, Shanghai 200233, China
 Email: bhan@ncgr.ac.cn
 Clone requests: bhan@ncgr.ac.cn
 This is rice cDNA est clone
 Web site: http://www.ncgr.ac.cn.
 FEATURES
 Location/Qualifiers
 1..718
 /organism="Oryza sativa"
 /mol_type="mRNA"
 /db_xref="taxon:4530"
 /clone="Y745ellp5"
 /clone_lib="Oryza sativa library (Han B)"
 ORIGIN
 Alignment Scores: 4.43e-89 Length: 718
 Pred. No.: 962.50 Matches: 198
 Score: 93.42% Conservative: 15
 Percent Similarity: 86.84% Mismatches: 15
 Best Local Similarity: 34.63% Indels: 3
 Query Match:

DB: 7 Gaps: 0

US-10-051-909-36 (1-553) x CR291514 (1-718)

QY 160 AlaGluTyrIleGlyArgGlySerLeuMetIleAlaAlaIleProAsnIleIleGly 179

DB 8 GCGAGCTACATGGCGCGAAGGGTCA-TTGATGATTGCTGCAATTCGCAATCATCATGGT 66

QY 180 TrpLeuAlaIleSerPheAlaAlaSerPheLeuTyrMetGlyArgGlyLeuGlu 199

DB 67 TGGCTTGCATCTCTTTGCAAGAGCATCATGTTCTTTATATGGACGATGCTCGAG 126

QY 200 GlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPro 219

DB 127 GGGTTGGTGTGGTGTCTATACGGTGGCGAGTTTACATAGCAGAAATATCACCT 186

QY 220 GlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePhe 239

DB 187 CAAACATGAGAGGCTCTTGGCTCAGTGAATCAGTATCTGTACCCGTTGGTATATTG 246

QY 240 LeuAlaTyrLeuLeuGlyMetPheIleProTyrPheLeuLeuAlaValIleGlyAlaLeu 259

DB 247 TTGGCATATTGCTCGGCATGTTGTTCTTGGAGGCTTCTGTCTGAATAGGAATCTTG 306

QY 260 ProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTyrLeuAla 279

DB 307 CTTGCACTGTGTGATACCTGGCTATTCTTCAATCCAGATCCCAAGATGGTTGGCA 366

QY 280 LysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThr 299

DB 367 AAGTGAACATGATGATGTTTGAGACTTCTTACAGTCTCTGAGGGATTTGAGACT 426

QY 300 AspileThrThrGluValAsnAspIleIleArgAlaValAlaSerSerSerLysArgThr 319

DB 427 GACATCAGCGCGGAAGTGAATATTAAGAGAGCAGTAGCTTCAGCAAAACAAAGGACA 486

QY 320 ThrIleSerPheGlnGluLeuAsnGlnLysIleTyrArgThrProLeuLeuGlyIle 339

DB 487 ACGATCGTTTCAAGAAATTAACAGAGAAATACCCGACCCCTTAATACATAGGAAT 546

QY 340 GlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSer 359

DB 547 GGCCTACTTGCTACTGCAACAGGTAAAGTGAATCACCAGAAATATTGTTTATGACGATG 606

QY 360 IlePheLysAlaIleGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIle 379

DB 607 ATCTTCAAGCAGCAGGTCTCACAAACAGTGACTTGGGTACATGTGCATGGGGGT-ATC 665

QY 380 GlnValLeuAlaThrGlyValThr 387

DB 666 CAGGTTCT-GCTACAGAGTTACA 688

RESULT 12

AJ794429

LOCUS

DEFINITION AJ794429 Antirrhinum majus whole plant Antirrhinum majus cDNA clone

018 3 05 d17, mRNA sequence.

ACCESSION AJ794429.1

VERSION

KEYWORDS

SOURCE

ORGANISM

Antirrhinum majus (snapdragon)

Antirrhinum majus

Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Plantaginaceae; Antirrhineae;

Antirrhinum

1 (bases 1 to 762)

Zachgo,S., Stueber,K., Saedler,H., Sommer,H. and Schwarz-Sommer,Z.

TITLE

Antirrhinum EST collection

UNPUBLISHED (2003)

Contact: Schwarz-Sommer Z

Molekulare Pflanzen-genetik

MPI fuer Zuechtungs-forschung

FEATURES

source

Carl-von-Linne Weg 10, D-50829, Germany.

1..762

/organism="Antirrhinum majus"

/mol_type="mRNA"

/db_xref="taxon:4151"

/clone="018_3_05.d17"

/tissue_type="whole plant"

/clone_lib="Antirrhinum majus whole plant"

ORIGIN

Alignment Scores:

Pred. No.: 4,87e-89 Length: 762

Score: 962.50 Matches: 189

Percent Similarity: 86.00% Conservative: 26

Best Local Similarity: 75.60% Mismatches: 34

Query Match: 34.63% Indels: 1

DB: 1 Gaps: 1

US-10-051-909-36 (1-553) x AJ794429 (1-762)

QY 88 SerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAla 107

DB 13 TCTTCTCAAGCTTTTAGGGACAGTCTCTTCTGTTGGCTTGTGTTAATTGGCT 72

QY 108 LeuGlyProIleGlnPheThrSerGlyPheSerProThrGlnAspAlaVal 127

DB 73 TTGGTCTCTATCCAGTTTGGTTTCAAGTGGTTATTCTTCCAGGACTCAAACCTGCTATT 132

QY 128 ValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnVal 147

DB 133 GTCAATGATCTTAAACTCACAGTTCACAGTTCGCGTATTGTTGTTCTTGTCCAGCTT 192

QY 148 GlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgGly 167

DB 193 GGAGCCATGGTTGGAGCAATAGCTAGTGTGTCAGATTGCTGAGTACATAGGGAGAAAGGG 252

QY 168 SerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLys 187

DB 253 TCTTTAATGATCGCTGGCCATACCTAATATCATTTGGTTGGCTTGGATTTCGTTGCCAGA 312

QY 188 AspAlaSerPheLeuTyrMetGlyArgLeuLeuGlyPheGlyValGlyIleIleSer 207

DB 313 GACATCTCATTTCTGTACATGGGAAGATTGTTGGAGATTGGGTCGGCATATCTCT 372

QY 208 TyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGly 227

DB 373 TATCTGTGCTGTATATATAGCTGAGATAGCTGAGATAGCTGAGATAGCTGAGATAGCTGAG 432

QY 228 SerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPhe 247

DB 433 TCAGTAAACAGCTCTCTGTACCATCGGATCATGCTAGCATATTATTCTTGGACTATT 492

QY 248 IleProThrArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGly 267

DB 493 TCGAATTGGAGATTGCTTGTGTTCTAGGAATATTGCTTGTCTGATATTGATACCTGGC 552

QY 268 LeuPheIleProGluSerProArgTyrPheLeuAlaLysMetAsnLeuThrGluAspCys 287

DB 553 CTCCTTTTTCATCCAGATCTCTTAGATGTTTGGCCAAATGGGGATGACAGAAAGATT 612

QY 288 GluThrSerLeuGlnValLeuArgGlyPheGluThrAspileThrThrGluValAsnAsp 307

DB 613 GAAGCTCTCTTCAAGTTCTTTCGGGGGTTTACGCTGATATTGAGTTGAATATCTGAG 672

QY 308 IleLysArgAlaValAlaSerSerLysArgThr---ThrIleSerPheGlnGluLeu 326

DB 673 ATAAGAAATCTGTGCTTCAACGAGCAGCGAGGAGCAATCCGCTTTGCCGATCTT 732

QY 327 AsnGlnLysLysTyrArgThrProLeuLeu 336

DB 733 AAGTTGAAAGATACTGTTTACCGTTGATG 762

```

RESULT 13
CD871121      701 bp      mRNA      linear      EST 11-JUL-2003
LOCUS        AZ02.117H06F010207 AZ02 triticum aestivum cDNA clone AZ02117H06,
DEFINITION  mRNA sequence.
ACCESSION   CD871121
VERSION     CD871121.1 GI:32554937
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
ORGANISM    Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
1 (bases 1 to 701)
Genoplane.
Genoplane, a major partnership french program in plant genomics
Unpublished (2003)
Contact: Genoplane
Genoplane
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplane' (http://www.genoplane.com
and http://genoplane-info.infobiogen.fr).
FEATURES
source
1..701
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="AZ02117H06"
/tissue_type="root"
/clone_lib="AZ02"

ORIGIN
Alignment Scores:
Pred. No.:      6.28e-86      Length:      701
Score:          932.00      Matches:     182
Percent Similarity: 90.48%      Conservative: 27
Best Local Similarity: 78.79%      Mismatches:  22
Query Match:     33.54%      Indels:      0
DB:              6          Gaps:          0

US-10-051-909-36 (1-553) x CD871121 (1-701)

Qy 285 GluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGlu 304
Db 7 GAGGACCTTGGAACTCTCTACAGTTCTGAGGGGATTTGAGACTGACATCACTCAGAA 66

Qy 305 ValAsnAspIleLysArgAlaValAlaSerSerLysArgThrThrIleSerPheGln 324
Db 67 GTGAATCATATAAAGAGAGAGAGTAAATATCAGCAACAAAGGGGGCGGATCCGTTCCAA 126

Qy 325 GluLeuAsnGlnLysLysTyArgThrProLeuLeuLeuGlyIleGlyLeuValLeu 344
Db 127 GAGTTAAACCAAGAAATTCGGATGCCCTGATCTAGAAATGGCCCTGTGTCTA 186

Qy 345 GlnAsnLeuSerGlyIleAsnGlyValLeuPheTyAlaSerSerIlePheLysAlaAla 364
Db 187 CAACAGCTAAGCGGAATCAACGCTACTACTGTCTATCGAAGTAGACTCTTCAAGCTGCA 246

Qy 365 GlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThr 384
Db 247 GGTATTACAAACAGTGAATTCGGCCACATCTGGACTTGGAGTATTCAGGTTCTGGCACT 306

Qy 385 GlyValThrThrTpLeuLeuAspArgAlaGlyArgIleLeuLeuIleIleSerThr 404
Db 307 CTAGTTACACCTGGTTACTAGACAGGGCTGGCGGGGTATCTCTACTCATATATCTTCT 366

Qy 405 SerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeuLysAspAsnIle 424
Db 367 GCTGGGATGACTATAAGCCTCTCTTGGCGTTGCCGTCATATATTTTATCAAGGACACTGTT 426

```

```

Qy 425 SerGlnAspSerAsnSerTyThrIleLeuThrMetIleSerLeuValGlyIleValSer 444
Db 427 TCACAAGACTCTCCACATATTACATATTGAGCATGGTCTCCTCTGCTATTGGCT 486

Qy 445 PheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIle 464
Db 487 TATGTTATGCCCTTCTCCTTCGGTATGGGGCCATTCATGGGTCTAATAATGTCGAGATT 546

Qy 465 LeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThr 484
Db 547 CTCCTCGGTGAGCATCAAGAGTCTCGCGGAGAGCTTCGGGACGCTGCCCAACTGGGTGACT 606

Qy 485 SerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPhe 504
Db 607 TCTTTGGATTAACGATGACGCAAACTTGCTGCTCAGCTGGAGTGTGGAGGTACATTT 666

Qy 505 LeuSerTyMetValValSerAlaPheThrIle 515
Db 667 GTGTCTTACATGCTCGTGAGCGCTTCACGCTC 699

RESULT 14
CNS0A8HU
LOCUS
DEFINITION
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTSL24ZD01 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION   BX819759
VERSION     BX819759.1 GI:42467369
KEYWORDS    HTC; GSLT_cDNA.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1675)
REFERENCE
1 Castelli V., Aury J.M., Jaillon O., Wincker P., Clepet C.,
Menard M., Cruaud C., Quetier F., Scarpeili C., Schachter V.,
Temple G., Caboche M., Weissenbach J. and Salanoubat M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
2 (bases 1 to 1675)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequençage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb?ggb?source=Arabidopsis.
FEATURES
source
1..1675
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strains="Col-0"
/db_xref="taxon:3702"
/clone="GSLTSL24ZD01"
/tissue_type="Adult vegetative tissue"
/plasmid="pCMVSPORR_6"
1..1675
/gene="At2g548020"

```

[illegible][illegible]

	295	LyMetAsnLeuThrGluAlaSpCyGluThrSerLeuGlnValLeuArgGlyPheIleuLys
QY	280	LyMetAsnLeuThrGluAlaSpCyGluThrSerLeuGlnValLeuArgGlyPheIleuLys
	741	AAATGGGGCGGTATCGGAGTTTGAAGCTGCACCTGAGGAAGCTCCGTGGGAAGAGCT
Db	800	AspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThr
QY	300	AspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThr
	801	GATATTCGGAGGAGCGCAGAGATCCAGGATTATATCGAAACTCTGGAAGGCTACCG
Db	860	GATATTCGGAGGAGCGCAGAGATCCAGGATTATATCGAAACTCTGGAAGGCTACCG

	Qy	320 ThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIle 339
		:::::.....:::
	Db	861 AAAGCCAGATGCTGGATTTCACAGAGAGATACATACGCCTGTCTTTATAGCTTTC 920
		:::.....:::
	Qy	340 GlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyTAlaSerSer 359
		:::.....:::
	Db	921 GGGTGTATGGTGTTTCAGCAGCTTTGAGGAGGAATCACGGAATAATGTCTTACACAGACTCG 980
		:::.....:::
	Qy	360 IlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIle 379
		:::.....:::
	Db	981 ATATTGAGCAAGCAGGTTTTCCC--ACAAGACTTGGCATGATATAATATGCTGCTCTT 1037
		:::.....:::
	Qy	380 GluValLeu-AlaThrGlyValThrTrpLeuLeuAspArgAlaGlyArgArgIleLe 399

Alignment Scores:

399 uLeuIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePh 419

Pred. No.: 3.88e-85 Length: 769
Score: 925.00 Matches: 177
Percent Similarity: 85.66% Conservative: 38
Best Local Similarity: 70.52% Mismatches: 36
Query Match: 33.29% Indels: 0
DB: 4 Gaps: 0

US-10-051-909-36 (1-553) x BI933646 (1-769)

```
Qy 181 LeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGly 200
Db 1 CTTTCCATCTGTTGCGAAGATCTGCTCTTCTTATACATGGGAAGATTATTGGAAGT 60

Qy 201 PheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerProGln 220
Db 61 TTTGGAGTCGCATATAATATCTTACACGGTTCCTGTATATATGCGGATAGCACCTCAG 120

Qy 221 AsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeu 240
Db 121 AACCTGAGAGGGGCCCTGGGCTCAGTTAACCCAGCTCTCTGTTCACATTGGGATCATGTTG 180

Qy 241 AlaTyrIleLeuGlyMetPheIleProTyrPheLeuAlaValIleGlyAlaLeuPro 260
Db 181 GCTTATTGTCTAGGACTTTTGTAAATTGGAGAGTGTCTGCTTTTCTTGGAAACATTGCC 240

Qy 261 CysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTyrPheLeuAlaLys 280
Db 241 TGCCTTGCAATTGATACCTGGGCTATTTTTCATCCAGAAATCTCTCGTGTGGTGGCCAA 300

Qy 281 MetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAsp 300
Db 301 ATGGGTCTGACAGATGATTTTGAACCTCTTTGCAAGTCTCCGAGGGTTCGATGTGAC 360

Qy 301 IleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerLysArgThrThr 320
Db 361 ATTTCCGTCGAAGTAATGAATTAAGAGGGCTGTAGCATCCACAGCCGAAAGTCAACA 420

Qy 321 IleSerPheGlnGluLeuAsnGlnLysIleTyrArgThrProLeuLeuLeuGlyIleGly 340
Db 421 ATACGTTTGCAGATCTCAACAAAGAGATATTGGCTGCTCTCATGATAGGCAATTGGA 480

Qy 341 LeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIle 360
Db 481 CTGCTTGTCTTACAACTCAGCGGAACCAATGGTGTGATCTTCTTATTCAGTAACATT 540

Qy 361 PheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGln 380
Db 541 TTCCTATCGCGCGGATTTCTCAAGTCATGCTCAACTTTAGGTTTGTGCTATCCAG 600

Qy 381 ValLeuAlaThrGlyValThrThrTripleLeuAspArgAlaGlyArgArgIleLeuLeu 400
Db 601 GTGGTGGCACTGCTGTTTCTACATGGCTGGTGGATAAACTGCGCGTAGGCTTTACTG 660

Qy 401 IleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValPhePheLeu 420
Db 661 ATTGCTCGTCAGTGAATGGCTGCTGCTCTTATTTGTTTCCATTGCAATTTCTTGGT 720

Qy 421 LysAspAsnIleSerGlnAspSerAsnSerTyr 431
Db 721 AAGGATTCGATGATGAGGATTTACCTTCTAT 753
```

Search completed: October 13, 2004, 19:46:51
Job time : 4511.99 secs

This Page Blank (uspto)

...is Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 11:33:18 ; Search time 62.1256 Seconds

(without alignments)
5121.591 Million cell updates/sec

Title: US-10-051-909-36

Perfect score: 2779

Sequence: 1 PSSSSSRPAGKKKKXNOG.....TIFVLSIQRLQWLPECL 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1801.5	64.8	501	2 Q9V609	Q9V609 oryza sativ
2	1757.5	63.2	490	2 Q9V608	Q9V608 beta vulgar
3	1746.5	62.8	488	2 Q9V608	Q9V608 arabidopsis
4	1734.5	62.4	487	2 Q9V608	Q9V608 arabidopsis
5	1713	61.6	515	2 Q9V608	Q9V608 arabidopsis
6	965.5	34.7	480	2 Q9V608	Q9V608 arabidopsis
7	951.5	34.2	463	2 Q9V608	Q9V608 arabidopsis
8	943.5	34.0	463	2 Q9V608	Q9V608 arabidopsis
9	939	33.8	482	2 Q9V608	Q9V608 arabidopsis
10	930	33.5	482	2 Q9V608	Q9V608 arabidopsis
11	920.5	33.1	496	2 Q9V608	Q9V608 arabidopsis
12	898.5	32.3	462	2 Q9V608	Q9V608 arabidopsis
13	890	32.0	463	2 Q9V608	Q9V608 arabidopsis
14	874.5	31.5	470	2 Q9V608	Q9V608 arabidopsis
15	868.5	31.3	478	2 Q9V608	Q9V608 arabidopsis
16	863	31.1	477	2 Q9V608	Q9V608 arabidopsis
17	853.5	30.7	474	2 Q9V608	Q9V608 arabidopsis
18	851.5	30.6	425	2 Q9V608	Q9V608 arabidopsis
19	843.5	30.4	458	2 Q9V608	Q9V608 arabidopsis
20	824	29.7	483	2 Q9V608	Q9V608 arabidopsis
21	822.5	29.6	479	2 Q9V608	Q9V608 arabidopsis
22	820.5	29.5	474	2 Q9V608	Q9V608 arabidopsis
23	808	29.1	454	2 Q9V608	Q9V608 arabidopsis
24	769	27.6	468	2 Q9V608	Q9V608 arabidopsis
25	768	27.6	804	2 Q9V608	Q9V608 arabidopsis
26	748.5	26.9	461	2 Q9V608	Q9V608 arabidopsis
27	737	26.5	498	2 Q9V608	Q9V608 arabidopsis
28	729	26.2	457	2 Q9V608	Q9V608 arabidopsis
29	698.5	25.1	497	2 Q9V608	Q9V608 arabidopsis
30	693.5	25.0	339	2 Q9V608	Q9V608 arabidopsis
31	691	24.9	482	2 Q9V608	Q9V608 arabidopsis

Q9V609 drosophila
AaQ23604 drosophila
Q9V608 drosophila
Q86P59 drosophila
Q7P1R5 anopheles g
Q04040 arabidopsis
Bac38338 mus muscu
Q9J1F3 mus muscu
Q9J1Z1 rattus norv
Q940K3 arabidopsis
Q8WU29 homo sapien
Q9NY64 homo sapien
Q6XUI2 bos taurus
Aap43920 bos tauru

ALIGNMENTS

RESULT 1

Q9V609 PRELIMINARY; PRT; 501 AA.
ID Q9V609
AC Q9V609
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE 10A191.3 protein.
GN Name=10A191.3;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
RN NCBI_TaxID=4530;
RX [1]
RP SOURCE FROM N.A.
RA Vyotskaia V.S., Schwartz J.R., Osborne B.I., Wing R., Yu G., Kwan A.,
RA Liu S., Lee J., Toriumi M., Luros J., Li J., Kremenetskaia I., Oji O.,
RA Theologis A.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL: AC007858; AAD39600.1; -;
DR Gramene: Q9V609;
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0005351; F: sugar porter activity; IEA.
DR GO: GO:0005215; F: transporter activity; IEA.
DR GO: GO:0008643; P: carbohydrate transport; IEA.
DR InterPro: IPR007114; MFS.
DR InterPro: IPR005828; Sub transporter.
DR InterPro: IPR003663; Sugar transporter.
DR InterPro: IPR005829; Sug transporter.
DR Pfam: PF00083; Sugar tr; 1.
DR PRINTS: PR00171; SUGTRNSPORT.
DR TIGRPFAMS: TIGR00879; SP; 1.
DR PROSITE: PS00850; MFS; 1.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
DR KW Sugar transporter; Transmembrane; Transporter.
SQ SEQUENCE 501 AA; 53291 MW; FBAB15CE2E769491 CRC64;

Query Match 64.8%; Score 1801.5; DB 2; Length 501;
Best Local Similarity 72.2%; Pred. No. 1.3e-111;
Matches 351; Conservative 62; Mismatches 64; Indels 9; Gaps 2;

QY 51 GEESGSDHNG-----VLRRLNTGSGWTMRSSRGSFAPGT---SSMAVLRESHVSAFL 101
Db 5 GEESGSDHNG-----VLRRLNTGSGWTMRSSRGSFAPGT---SSMAVLRESHVSAFL 64
QY 102 CTLIVAGLPIQFGTSGFSSPTQDAMVRDLNLISIEFSAFGSLSNVGMVGAIASQMAE 161
Db 65 CTLIVAGLPIQFGTSGFSSPTQDAMVRDLNLISIEFSAFGSLSNVGMVGAIASQMAE 124
QY 162 YIGRKGSGLMAAIPNTIGWLAI5FAKDA5FLYNGRLLEGGVIGIISYTPVPIAEISPN 221

```

Db 125 YIGKGLMTAAIPNIIIGWLAISPAKDSFLFMGRLLLEGFGVGVISVVPVYIAEAPQT 184
QY 222 MRGALGSVNQSVTFGIFLAYLLGMFIPWRLLAVIGALPCTMLIPGLFFIPESPRWLAKM 281
Db 185 MRGALGSVNQSVTFGIFLAYLLGMFIPWRLLAVIGALPCTMLIPGLFFIPESPRWLAKM 244
QY 282 NLTEDETSQVLRGFTDITTEVNDIKRAVASSKRTTISFOELNOKKRTPLILGIGL 341
Db 245 GKMEDFESSQVLRGFTDITTEVNDIKRAVASSKRTTISFOELNOKKRTPLILGIGL 304
QY 342 LVLQNLGSLGVLYASSIFKAAAGVTNSDLATCSLGAIQVLTGVTWLLDRAGRILLI 401
Db 305 LVLQNLGSLGVLYASSIFKAAAGVTNSDLATCSLGAIQVLTGVTWLLDRAGRILLI 364
QY 402 ISTSGMTCLLAVSVFLLKDNISQDSNYVILTMISLGVISFVTFPSFGMGAIPWLMM 461
Db 365 ISTSGMTITLVVSVSPFVKDNITNGSLYVMSMLSLVGLVAFVIFSGLGAIPWIIM 424
QY 462 SEILPVSTKSLGSGIATLANWLTSAITMTNMLTWSVGGTFLSYMVVSFAFTVFWLW 521
Db 425 SEILPVSTKSLGSGIATLANWLTSAITMTNMLTWSVGGTFLSYMVVSFAFTVFWLW 484
QY 522 VPETKG 527
Db 485 VPETKG 490

RESULT 2
Q39416 PRELIMINARY; PRT; 490 AA.
AC Q39416;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-2004 (TrEMBLrel. 26, Last annotation update)
DE Integral membrane protein.
OS Beta vulgaris (Sugar beet).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Amaranthaceae; Beta.
OX NCBI_TaxID=161934;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96351183; PubMed=8742332;
RA Chou T.J., Bush D.R.;
RT "Molecular cloning, immunochemical localization to the vacuole, and
RT expression in transgenic yeast and tobacco of a putative sugar
RT transporter from sugar beet."
RL Plant Physiol. 110:511-520(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Chou T.-J., Bush D.R.;
RA Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RL SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL; U43629; AB53155.1; -.
DR PIR; T14545; T14545.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0005351; F: sugar porter activity; IEA.
DR GO; GO:0005215; F: transporter activity; IEA.
DR GO; GO:0008643; P: carbohydrate transport; IEA.
DR InterPro; IPR007114; MPS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; Sugar_tr; 1.
DR PRINTS; PR00171; SUGETRNSPORT.
DR TIGERFAMS; TIGR00879; SP; 1.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
DR Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 490 AA; 53020 MW; 3F8C96D26989CB8C CRC64;

```

```

Query Match 63.2%; Score 1757.5; DB 2; Length 490;
Best Local Similarity 66.2%; Pred. No. 1.1e-108;
Matches 337; Conservative 76; Mismatches 73; Indels 23; Gaps 3;

QY 39 MGGGSRNRGAGAGEBSGDHGVLRRLPLNTGSRWYRMSRSSQSPAGTSSMAVLRSHVS 98
Db 1 MSSDEAGLGGG-----GDLRKPFLHTGSMVRMGSRSSL--MGSSQVIRESSIS 49
QY 99 AFLCTLIIVALOPTQGFSTGSSSTQDAMVDNLNLSIFSAFGLSNGVMGVAIASGQ 158
Db 50 VLACVLIIVALOPTQGFSTGSSSTQSAITNELGLSVAEYSWFGLSNGVMGVAIASGQ 109
QY 159 VAEYIGRKGSLMIAIPNIIIGWLAISPAKDSFLFMGRLLLEGFGVGVISVVPVYIA 218
Db 110 ISEYIGRKGSLMIAIPNIIIGWLAISPAKDSFLFMGRLLLEGFGVGVISVVPVY 169
QY 219 PQNMRGALGSVNQSVTFGIFLAYLLGMFIPWRLLAVIGALPCTMLIPGLFFIPES 278
Db 170 PQNMRGALGSVNQSVTFGIFLAYLLGMFIPWRLLAVIGALPCTMLIPGLFFIPES 229
QY 279 AKNNLTEDCETSLQVLRGFTDITTEVNDIKRAVASSKRTTISFOELNOKKRTPL 338
Db 230 AKNGMVEEFVSLQVLRGFTDITTEVNDIKRAVASSKRTTISFOELNOKKRTPL 289
QY 339 IGLLVQLNLSGIVGVLYASSIFKAAAGVTNSDLATCSLGAIQVLTGVTWLLDRAG 398
Db 290 NGLLILQLLSGIVGVLYASSIFKAAAGVTNSDLATCSLGAIQVLTGVTWLLDRAG 349
QY 399 LLIITSGMTCLLAVSVFLLKDNISQDSNYVILTMISLGVISFVTFPSFGMGAIP 458
Db 350 LLIIVSSGMTCLLAVSVFLLKDNISQDSNYVILTMISLGVISFVTFPSFGMGAIP 409
QY 459 LMMSEILPVSTKSLGSGIATLANWLTSAITMTNMLTWSVGGTFLSYMVVSFAFT 518
Db 410 IIMSEILPVSTKSLGSGIATLANWLTSAITMTNMLTWSVGGTFLSYMVVSFAFT 469
QY 519 VLWVPETKXNSRGDTIFVSLISQRLQW 547
Db 470 VLWVPETKGR-----LEEIQW 486

RESULT 3
Q93YP9 PRELIMINARY; PRT; 488 AA.
ID Q93YP9;
AC Q93YP9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Similar to integral membrane protein.
GN Name=Atg1919450;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Zaker J., Theologis A., Davis R.W.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Zaker J., Theologis A., Davis R.W.;

```


Db 8 EEARND---LRPFHTGWSYRMSRQSSM---MGSSQVIRDSSISVLACVLIVLALGPI 60

Qy 112 QFGTSGFSSPTQAMVDRDLNISERFASGSLSNVGGVGAIAAGOMAEYIGRKSLMI 171

Db 61 QFGTCGYSSPTQAAITKDLGTLTVEYSVFGSLSNVGGVGAIAAGOMAEYIGRKSLMI 120

Qy 172 AAINIIGLWLAISPAKDAFLYMRLLGEGVGIIISYTPVYIAETSPQNRGALGSVNO 231

Db 121 AAINIIGLWLAISPAKDAFLYMRLLGEGVGIIISYTPVYIAETSPQNRGALGSVNO 180

Qy 232 LSVTFGFLAYLLQMFIPWLLAVIGALPCTMLIPGLFFIPESPRMLAKNLTDETCSL 291

Db 181 LSVTIGIMLAVLLGLFVPRILAVILGILPCTLLIPGLFFIPESPRMLAKNLTDETCSL 240

Qy 292 QVLRGFTDITTEVNDIKRAVASSKETTISFQELNCKKYRTPLLGILGLLVQNLSGIN 351

Db 241 QVLRGFTDITTEVNEIKRVAASSTKENTRVFDLKRRTYFFLWVGIGLLVQLQGIN 300

Qy 352 GVLRYASSIIFKAAGVTNSDLATCSLGAIOVLATGVITWLLDRAGRILLIISTSGMTLCL 411

Db 301 GVLRYSSITFESAGVTSSNAATFGVGAIOVATAISTWLVDKAGRLLLTISVGMNTISL 360

Qy 412 LAVSVFPLKDNISQDSNYIILTMISLGVISFVTFSGMGAIPKLMSEILPYSIKS 471

Db 361 VIVAAAFYLFKEFVSPDSMSYLSVGVGVWVAMVFFSLGMPPLWIMSEILPVNIK 420

Qy 472 LGGSIATLANWLTSTFAITMTNMLTWSVGTFTLSYMWVSAFTIVFVLWVPETKG 527

Db 421 LAGSIATLANWFFSWLITMTANLLANSGGTFLYGLVCAFTVFFVLWVPETKG 476

RESULT 5

Q9LN48 PRELIMINARY; PRT; 515 AA.

AC Q9LN48

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)

DE F18014.22.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S., Kim C., Altafi H., Bei Q., Chin C., Chio J., Choi E., Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A., Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A., Ecker J.R.;

RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S., Kim C., Altafi H., Bei Q., Chin C., Chio J., Choi E., Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A., Toriumi M., Vaysberg M., Yu G., Federspiel N., Theologis A., Ecker J.;

RA Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C., Shinn P., Altafi H., Bei Q., Chin C., Chio J., Choi E., Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N., Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A., Ecker J.;

RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

CC -!- SURCELLULAR LOCATION: Integral membrane protein (By similarity).

CC -!- SIMILARITY: Belongs to the sugar transporter family.

DR EMBL; AC035808; AAF79445.1; --. Integral to membrane; IEA.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0005351; F:sugar porter activity; IEA.

DR GO; GO:0005215; F:transporter activity; IEA.

DR GO; GO:0008643; P:carbohydrate transport; IEA.

DR InterPro; IPR007114; MFS.

DR InterPro; IPR005828; Sub transporter.

DR InterPro; IPR003663; Sugar_transp.

DR InterPro; IPR005829; Sug_transp.

DR Pfam; PF00883; Sugar_tr; 1.

DR PRINTS; PRO0171; SUGTRNSPORT.

DR TIGRFAMs; TIGR00879; SP; 1.

DR PROSITE; PS00850; MFS; 1.

DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.

DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.

KW Sugar transport; Transmembrane; transport.

SK SEQUENCE 515 AA; 55930 MW; 80574480B64EF8D4 CRC64;

Query Match

Best Local Similarity 66.4%; Score 1713; DB 2; Length 515;

Matches 334; Conservative 67; Mismatches 68; Indels 34; Gaps 4;

Qy 52 EESGDHGVLRRLPNTGWSYRMSRQSSPAGTSSVAVLRSHVSAFLCTLIVLALGPI 111

Db 9 EEGEND---LRPFHTGWSYRMSRQSSM---LESSQVIRDSSISVLACVLIVLALGPI 61

Qy 112 QFGTSGFSSPTQAMVDRDLNISERFASGSLSNVGGVGAIAAGOMAEYIGRKSLMI 171

Db 62 QFGTCGYSSPTQAAITKDLGTLTVEYSVFGSLSNVGGVGAIAAGOMAEYIGRKSLMI 121

Qy 172 AAINIIGLWLAISPAKDAFLYMRLLGEGVGIIISYTPVYIAETSPQNRGALGSVNO 231

Db 122 AAINIIGLWLAISPAKDAFLYMRLLGEGVGIIISYTPVYIAETSPQNRGALGSVNO 181

Qy 232 LSVTFGFLAYLLQMFIPWLLAVIGALPCTMLIPGLFFIPESPRMLAKNLTDETCSL 291

Db 182 LSVTIGIMLAVLLGLFVPRILAVILGILPCTLLIPGLFFIPESPRMLAKNLTDETCSL 241

Qy 278 ---LAKNLTDETCSLQVLRGFTDITTEVNDIKRAVASSKETTISFQELNCKKYRTPLLGILGLLVQNLSGIN 351

Db 242 VALLAKMGLTDDFETSLQVLRGFTDITTEVNEIKRVAASSTKENTRVFDLKRRTYFFLWVGIGLLVQLQGIN 301

Qy 325 ELNQKYFTPLLGILGLVQLNLSGIVGLFVASSIFKAAGVTNSDLATCSLGAIOVLAT 384

Db 302 DLKRRYFFPLMVGIGLLALQQLGGINGVLFYSSTIFESAGVTSSNAATFGVGVVQVAT 361

Qy 385 GYTTWLLDRAGRILLIISTGWTICLLAVSVFPLKDNISQDSNYIILTMISLGVIS 444

Db 362 GIATWLVDKAGRLLMISGIMTISLIVAVAFVLFKEFVSPDSNMVILSWSVGVVA 421

Qy 445 FVITFSGMGAIPWLMSEILPVSIKSLGGSIIATLANWLTSTFAITMTNMLTWSVGTFTLS 504

Db 422 MVISCSLGMGPPLWIMSEILPVNIKLAGSIATLANWLVSWLVTMTANLLANSGGTFL 481

Qy 505 LSWMVSAFTIVFVLWVPETKG 527

Db 482 TLYALVCGFTVVVFLWVPETKG 504

RESULT 6

Q9LEG2 PRELIMINARY; PRT; 480 AA.

AC Q9LEG2

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Putative sugar transporter.

GN Name=st3;

OS Lycopersicon esculentum (Tomato).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamnids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
R2 SEQUENCE FROM N.A.
RC T18SUB=Root;
RA Ferrol N., Pozo M.J., Barea J.M., Azcon-Aguilar C.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL: AJ278765; CAC00637.2; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0005351; F:sugar porter activity; IEA.
DR GO: GO:0005215; F:transporter activity; IEA.
DR GO: GO:0008643; P:carbohydrate transport; IEA.
DR InterPro: IPR007114; MFS.
DR InterPro: IPR005828; Sub_transporter.
DR InterPro: IPR003663; Sugar_transp.
DR InterPro: IPR005829; Sug_transporter.
DR Pfam: PF00083; Sugar_tr; 1.
DR TIGRFAMs: TIGR00879; SP; 1.
DR PROSITE: PS00850; MFS; 1.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 480 AA; 52009 MW; 1B307EAA7C016D2F CRC64;

Query Match 34.7%; Score 965.5; DB 2; Length 480;
Best Local Similarity 42.1%; Pred. No. 5.2e-56;
Matches 183; Conservative 108; Mismatches 139; Indels 5; Gaps 2;

QY 93 RESHVSFLCTLIIVAGPIQFGTSGFSSPTQDAMVRDNLNLSIEFSAFGLSNVGVWG 152
Db 38 QDRCWYSLTVAVRGYSGLSCAGYSSPTQSAIREDLNLSIAQISLFGSIWTFGAMIG 97
QY 153 AIASGQMAEYIGRKGSLMIAIPNIIGWLAIISFAKDAFLYMGRLLEFGVGVIISYTPV 212
Db 98 AITSGPIADYIGRKGAMSSGFCVAGWLAIFFAQGALALDIGLATGYGMGVSYVVPV 157
QY 213 YIABISQNRGALGSVNLQSVTFGIFLAYLLGMFIPWRLLAVIGALPCTNMLIPGLFFIP 272
Db 158 FIABIAPKDLRGALTTINQLMICGVSVFIIGTMTWRTALTGLIPCAILLFGLFIIP 217
QY 273 ESPRLAKMLTDECTSLQVLRGFETDITTEVNDIKRAVASSKRTTISFQELNOKKYR 332
Db 218 ESPRLAKIGHQEFELAKRLKRDADISEAAEIKDYETLEKLPKNLFDLPQRYS 277
QY 333 TPLLGLGLVQLNSGVLFYASSIFKAAGVTNSDLATCSLGAIVLATGVTTWLLD 392
Db 278 SSLIVGVLMVFPQFGGICFTGSIFFESSGP-SSDIGTIIYAIQVPIALGAALID 336
QY 393 RAGRIILLISTSGMTLCLLAVSVVFPFLKONISQDSNSYILTMISLVGIVSVFVTFSG 452
Db 337 RTGRKPLLVSGLTVIGCILTGISFYMKGH---EMAIKAAPILAVTGILVYIGSFVSG 392
QY 453 MGAIPWLMSEILPVSIKSGGSIATLANMLTSPAITMTNMLTWSVGGTFLSYVMVSA 512
Db 393 MGAVPVVMSEIYINIKGAAGSLATLVNFWGAWACSYTFNFMWNSTFGTFLVLYAVNA 452
QY 513 FTIVFVVLWVPETKG 527
Db 453 LSILFVFKIVPETKG 467

RESULT 7
Q8LFR8 PRELIMINARY; PRT; 463 AA.
AC Q8LFR8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative sugar transporter.
DE Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
R2 SEQUENCE FROM N.A.
RC MEDLINE=2208475; PubMed=12093376;
RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation."; 3:RESEARCH0029-RESEARCH0029(2002).
RN [2]
R2 SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL: AY084684; AAM61246.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0005351; F:sugar porter activity; IEA.
DR GO: GO:0005215; F:transporter activity; IEA.
DR GO: GO:0008643; P:carbohydrate transport; IEA.
DR InterPro: IPR007114; MFS.
DR InterPro: IPR005828; Sub_transporter.
DR InterPro: IPR003663; Sugar_transp.
DR InterPro: IPR005829; Sug_transporter.
DR Pfam: PF00083; Sugar_tr; 1.
DR PRINTS: PR00171; SUGTRNSPORT.
DR TIGRFAMs: TIGR00879; SP; 1.
DR PROSITE: PS00850; MFS; 1.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 463 AA; 49695 MW; 2ED3D4ED040EC2A7 CRC64;

Query Match 34.2%; Score 951.5; DB 2; Length 463;
Best Local Similarity 42.3%; Pred. No. 4.3e-55;
Matches 191; Conservative 102; Mismatches 140; Indels 5; Gaps 2;

QY 100 FICTLIIVAGPIQFGTSGFSSPTQDAMVRDNLNLSIEFSAFGLSNVGVWGAIASGQM 159
Db 29 YLSTFVAVGPPAFGSCAGYSSPAQAARNDLSLTIAEFLSFLGSLITFCAMIGAITSGPI 88
QY 160 ABEYIGRKGSLMTAAIPNIIGWLAIISFAKDAFLYMGRLLEFGVGVIISYTPVYAEISP 219
Db 89 ADLVGRKGMVRVSSAPFCVVGWLAIIFAKGVVALDGLRATGVMGAFSVVVFIFAEIAP 148
QY 220 QNMGALGSVNLQSVTFGIFLAYLLGMFIPWRLLAVIGALPCTNMLIPGLFFIPESPRWLA 279
Db 149 KTFRGALITLNLQILICTGVSVSFIIGTLVTRVLALIGIIPCAASFLGLFPIESPRWLA 208
QY 280 KMLTDECTSLQVLRGFETDITTEVNDIKRAVASSKRTTISFQELNOKKYRTPELLGI 339
Db 209 KYGRDTEFEAALRKLKRGKADISEAAEIQDYETLERLPKAKMLDLFQRYIRSVLIAF 268
QY 340 GLLVLQNLSGINGVLFYASSIFKAAGVTNSDLATCSLGAIOVLATGVTTWLLDRAGRRL 399
Db 269 GLMVFQQGGINGICFYTSSIFEQAGFP-TRLMIIYAVLQVITALNAPIVDRAGRPL 327
QY 400 LIISTSGMTLCLLAVSVVFFLKONISQDSNSYILTMISLVGIVSVFVTFPSGMAIPWL 459
Db 328 LLVSATGLVIGCLIAAVSFYLVKH---DMAEAHPVLAVVGMVYIGSFSGAGMAMPVW 383
QY 460 VMSEILPVSIKSGGSIATLANMLTSPAITMTNMLTWSVGGTFLSYVMVSAFTVVFV 519
Db 384 VMSEIFPINIKGVAGGMATLVNFWGAWAVSYTFNFMWNSSYGTFLIYAINALAIIVFI 443
QY 520 LWVPETKG 527
Db 444 AIVPETKG 451
```

QY	160	ABYIGRKGSLMIAAIPNIIIGWLAISPAKADASFLYKMGRLLEGFGVGIISYTVPVYIAISP	219
DB	89	ADLVGRKGMEVSSAFVGVGLWLAIFAKGVVVALDGLRATGYGNGAFSYYVPIAIAIP	148
QY	220	QNMESGLSVNQLSVTGTGIFLAVILGMFIPWRLAVICALPCTMLIPGLFFIPESPRWLA	279
DB	149	KTFKALTTLLNQIILICTGVSFIIIGLTVRWVLAIGIIPCAASFGLGLFFIPESPRWLA	208
QY	280	KYNLTDCETSQVLRGPETDITTEVNDIKRAVASSKRTTISFOELNQKRYRTPLLLGI	339
DB	209	KVGRDTEFAALRLRGKKADISEAAAIQDIETLERLPKAKMLDLFQRRYRSVLIAF	268
QY	340	GLLVONLSGNGVLFYASSTFKAGVNSDLATCSLGAIQVLTATVTTMLDLAGRRL	399
DB	269	GLMVQFQGGINGICFYTSSIFEQAGPP-TRLGMIYAVLVQVITALNAPIVDRGRKPL	327
QY	400	LIISTSGMTCLLAVSVVFFFLKDNISQDSNSYITLMSLVGIVSVFVITFSFGMAIPWL	459
DB	328	LLVSATGLVIGCLLAASVFLKVH---DMAEAVPVLAUVGIMVYICGSFAGMGMPWV	383
QY	460	MMSEILPVSIKSLGSGTATLANWLTSPAITMTTMLTWSVGTFLSVVVSATFIVV	519
DB	384	VNSEIFPINIKGVAGMATLVNWFAGMAVSYTFNLSWSYGTFLIYAALNALIVFI	443
QY	520	LWVPETKG 527	
DB	444	AIVPETKG 451	
RESULT 9			
ID	Q8VZ15	PRELIMINARY;	PRT; 482 AA.
AC	Q8VZ15		
DT	01-MAR-2002	(TrEMBLrel. 20, Created)	
DT	01-MAR-2002	(TrEMBLrel. 20, Last sequence update)	
DT	05-JUL-2004	(TrEMBLrel. 27, Last annotation update)	
DE	AT5G18840/F17K4.90		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.		
OX	NCBI_TaxID=3702;		
RP	SEQUENCE FROM N.A.		
RA	Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,		
RA	Banh J., Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,		
RA	Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,		
RA	Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,		
RA	Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,		
RA	Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,		
RA	Yamamura Y., Yu G., Yu S., Shinzaki K., Davis R.W., Theologis A.,		
RA	Ecker J.R.;		
RL	Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,		
RA	Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,		
RA	Hsuan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,		
RA	Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,		
RA	Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,		
RA	Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,		
RA	Yamada K., Yu G., Yuan S., Shinzaki K., Davis R.W., Theologis A.,		
RA	Ecker J.R.;		
RL	Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.		
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (by similarity).		
CC	-1- SIMILARITY: Belongs to the sugar transporter family.		
DR	EMBL; AY064144; AAL36051.1; -		
DR	EMBL; BT000608; AAN18269.1; -		
DR	GO; GO:0005351; F:sugar porter activity; IEA.		
DR	GO; GO:0005215; F:transporter activity; IEA.		
DR	GO; GO:0008643; P:carbohydrate transport; IEA.		
Query Match 34.0%; Score 943.5; DB 2; Length 463;			
Best Local Similarity 42.1%; Pred. No. 1.5e-54;			
Matches 180; Conservative 102; Mismatches 141; Indels 5; Gaps 2;			
QY	100	FLCTLVALGPIQGTSGFSPTQDAMVRDLNLSIEFGAFGLSVNMGVGAISGQM	159
DB	29	YLSIFVAVCGSFAGSCAGYSSPAQAIRNDLSITIAEFLSGLLTFGAMIGAITSGPI	88

```
DR InterPro: IPR007114; MFS.
DR InterPro: IPR005828; Sub_transporter.
DR InterPro: IPR003663; Sugar_transp.
DR Pfam: PF00083; Sugar tr; 1.
DR PRINTS: PRO0171; SUGTRNSPORT.
DR TIGRFAMs: TIGR00879; SP; 1.
DR PROSITE: PS00850; MFS; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 482 AA; 5192 MW; 24E78EB630F8BFA0 CRC64;

Query Match 33.8%; Score 939; DB 2; Length 482;
Best Local Similarity 42.0%; Pred. No. 3e-54;
Matches 183; Conservative 91; Mismatches 154; Indels 8; Gaps 2;

QY 94 ESHVSAFCLTILVALGPIQFGTSGFSPSTQDAMVRDLNLSISFSAFSLSNVGMVGA 153
DB 39 ESYLWVLFSTFVAVCGSFPGSCVGSAPTQSIRQDLNLSLAFSMPGSLITIGMLGA 98
QY 154 IASGQMAEYIGRKSLMAAIPNIIGWLAIISFAKDAFLYMGRLLEGFGVGIISYTPVY 213
DB 99 VMSGKISDFSGRKGAMRTSACFCITGWLAVFTKGALLLDVGRFTGYGIVFVYVY 158
QY 214 IAEISPNMRGALGSVNQSLVTFGIFLAYLGMIPWRLLAIVIGALPCTMLIPGLFFIPE 273
DB 159 IAEISPNKRLGGLTTLNQLMTIVIGSSVFLIGSLISWKTLALTGLAPCVILLFGLCFIPE 218
QY 274 SPRWLAKMNLTEDETSQVLRGPETDITTEVNDIKRAVASSSKRTTISFOELNOKKYRT 333
DB 219 SPRWLAKACHEKFRVALQKRGKADITNEADGIQVSIQALEILPKARIQDLVSKYGR 278
QY 334 PLLLGIGLLVLQNLGSLGVLFYASSIFKAAGVTNSDLATCSLGAIOVLATGVTWLLDR 393
DB 279 SVIIGVSLWVQFVGINGIGFYASFTFKAGFTSGKLGITAIACVQVPIVLGLTILDK 338
QY 394 AGRILLIISTGMLTCLLAVSVVFLKDNISQDSNSYIITMISLV--GIVSFVITFSF 451
DB 339 SGRPLIMISAGIFLGCILTGTFTLLK-----GQSLLEWVPSLAVGGVLIYVAAPSI 392
QY 452 GMGAIPWLMWSEILPVSIKSLGSGIATLANLTLTSPATMTNMLTWSVGTFSLYMWVS 511
DB 393 GMGPVPWIMSEIPIPVKAGISLVLVNWSGAWAVSYTFNFLMSWSSPGTFYLYSAFA 452
QY 512 AFTIVFVVLWVPETKG 527
DB 453 AATIFVAKWVPETKG 468

RESULT 10
Q8LB19 PRELIMINARY; PRT; 482 AA.
AC Q8LB19;
DT 01-OCT-2002 (TREMELrel. 22, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DE 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Sugar transporter-like protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2208475; PubMed=12093376;
RA Haas B.J., Volkovskiy N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation."
RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
```

```
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL: AY087180; AAM64736.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0005351; F:sugar porter activity; IEA.
DR GO: GO:0005215; F:transporter activity; IEA.
DR GO: GO:0008543; P:carbohydrate transport; IEA.
DR InterPro: IPR007114; MFS.
DR InterPro: IPR005828; Sub_transporter.
DR InterPro: IPR003663; Sugar_transp.
DR Pfam: PF00083; Sugar tr; 1.
DR PRINTS: PRO0171; SUGTRNSPORT.
DR TIGRFAMs: TIGR00879; SP; 1.
DR PROSITE: PS00850; MFS; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 482 AA; 52034 MW; 20B39BF33BF80F10 CRC64;

Query Match 33.5%; Score 930; DB 2; Length 482;
Best Local Similarity 41.7%; Pred. No. 1.2e-53;
Matches 182; Conservative 91; Mismatches 155; Indels 8; Gaps 2;

QY 94 ESHVSAFCLTILVALGPIQFGTSGFSPSTQDAMVRDLNLSISFSAFSLSNVGMVGA 153
DB 39 ESYLWVLFSTFVAVCGSFPGSCVGSAPTQSIRQDLNLSLAFSMPGSLITIGMLGA 98
QY 154 IASGQMAEYIGRKSLMAAIPNIIGWLAIISFAKDAFLYMGRLLEGFGVGIISYTPVY 213
DB 99 VMSGKISDFSGRKGAMRTSACFCITGWLAVFTKGALLLDVGRFTGYGIVFVYVY 158
QY 214 IAEISPNMRGALGSVNQSLVTFGIFLAYLGMIPWRLLAIVIGALPCTMLIPGLFFIPE 273
DB 159 IAEISPNKRLGGLTTLNQLMTIVIGSSVFLIGSLISWKTLALTGLAPCVILLFGLCFIPE 218
QY 274 SPRWLAKMNLTEDETSQVLRGPETDITTEVNDIKRAVASSSKRTTISFOELNOKKYRT 333
DB 219 SPRWLAKACHEKFRVALQKRGKADITNEADGIQVSIQALEILPKARIQDLVSKYGR 278
QY 334 PLLLGIGLLVLQNLGSLGVLFYASSIFKAAGVTNSDLATCSLGAIOVLATGVTWLLDR 393
DB 279 SVIIGVSLWVQFVGINGIGFYASFTFKAGFTSGKLGITAIACVQVPIVLGLTILDK 338
QY 394 AGRILLIISTGMLTCLLAVSVVFLKDNISQDSNSYIITMISLV--GIVSFVITFSF 451
DB 339 SGRPLIMISAGIFLGCILTGTFTLLK-----GQSLLEWVPSLAVGGVLIYVAAPSI 392
QY 452 GMGAIPWLMWSEILPVSIKSLGSGIATLANLTLTSPATMTNMLTWSVGTFSLYMWVS 511
DB 393 GMGPVPWIMSEIPIPVKAGISLVLVNWSGAWAVSYTFNFLMSWSSPGTFYLYSAFA 452
QY 512 AFTIVFVVLWVPETKG 527
DB 453 AATIFVAKWVPETKG 468

RESULT 11
O65799 PRELIMINARY; PRT; 496 AA.
ID O65799;
AC O65799;
DT 01-AUG-1998 (TREMELrel. 07, Created)
DT 01-AUG-1998 (TREMELrel. 07, Last sequence update)
DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)
DE ERD6 protein (Putative sugar transporter) (Putative zinc finger
DE protein AT2F1).
DE Names:SUGTL1; Synonym=At1G08930;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98213606; PubMed=9545564;
RA Kiyosue T., Abe H., Yamaguchi-Shinozaki K., Shinozaki K.;
```

"ERD6, a cDNA clone for an early dehydration-induced gene of Arabidopsis, encodes a putative sugar transporter.",
 RN Biochim. Biophys. Acta 1370:187-191 (1998).
 (2)
 RA Gy I., Kreis M., Lechamy A.;
 RP Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RL
 RP SEQUENCE FROM N.A.
 RP Yamada K., Dale J.M., Hsuan V.W., Onodera C.S., Quach H., Toriumi M., Wong C., Wu H.C., Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J., Lam B., Lin J., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RP Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -!- SIMILARITY: Belongs to the sugar transporter family.
 DR EMBL; D89051; BAA25989.1; -.
 DR EMBL; AJ249967; CAB64732.1; -.
 DR EMBL; BT008661; AAP40473.1; -.
 DR PIR; T52132; T52132.
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR GO; GO:0005351; F: sugar porter activity; IEA.
 DR GO; GO:0005215; F: transporter activity; IEA.
 DR GO; GO:0008643; P: carbohydrate transport; IEA.
 DR InterPro; IPR000005; HTHAtaC.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub transporter.
 DR InterPro; IPR003663; Sugar_transpt.
 DR InterPro; IPR005829; Sug_transporter.
 DR Pfam; PF00083; Sugar_tr; 1.
 DR PRINTS; PR00171; SUGTRANSPORT.
 DR TIGRFS; TIGR00879; SP; 1.
 DR PROSITE; PS00041; HTH_ARAC_FAMILY_1; UNKNOWN_1.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 DR Sugar transport; Transmembrane, transport.
 KW SEQUENCE 496 AA; 54354 MW; 25CC9D21ADCI55CC CRC64;
 SQ
 Query Match 33.1%; Score 920.5; DB 2; Length 496;
 Best Local Similarity 41.4%; Pred. No. 53e-53;
 Matches 186; Conservative 91; Mismatches 159; Indels 13; Gaps 3;
 QY 98 SAFLCTLIVAGPTQFGTSGFSPTQDMVRDLNLSISFSAFGSLSNVGMVGAISG 157
 DB 57 SVFLSTFVAVSGSFTGCGVGFSSGACAGITKDLSLVAEYSMEGSLTLTGLGAVFSG 116
 QY 158 QMAEYIGRKGSLMTAAIPNIGLWLAISPAKDALFLYNGRLLGEGVGIIIVTPVYIAEI 217
 DB 117 KVADVLGRKRLTFCBFCITGWLQVALQANWMDGCRLLGLGIVGIFSVIPVYIAEI 176
 QY 218 SPQNRGALGVNQLSVTFGIFLAYLLGMFIPWRLAVIGALPCTMLIPGLFFIPSPRW 277
 DB 177 APKVRGSGFVFNQMLQMGCSLFLIIGNFIPWRLITVGLVPCVFHVFCLFFIPSPRW 236
 QY 278 LAKNMLTDCSTLQVLEGFTDITTEVNDIKRAVASSKETTISFQELNKKYKRTPLL 337
 DB 237 LAKGRDKECKSSQRLGSGSDISREANTIRDTDMTENGGETYKMSLFQRRYAYPLII 296
 QY 338 GIGLVLQNLGSLGVLPYASSIKFAAGVTNLSDLATCSLGAIOVLATGVTTWLLDRAGR 397
 DB 297 GVLGMLFQOLCGSGGVTYASSLNFKGGFP-SAIGTSVIATIMVPKAMLATVLVDKMGRR 355
 QY 398 ILLIISTGMLCLLANSVVFELKDNISQDSNSYIIL-----TMSLGVISFVITPFGM 453
 DB 356 TLLMASCANGSLALLSV-----SYGQSGFILPELTPITFCIGVLGHVIFPANGM 407
 QY 454 GAIPMLMSEILPVSIKSLGSGSIATLANWLTSPFAITMTNLMWTSVGGTFLSNVYSAF 513
 DB 408 GGLPFIWIAEIPMNKVSAGTLVTVTNWLFGLWIIITYTFNMLEWNASGMFLIFSVMAS 467
 QY 514 TIVFVVLWVPETKGNXSGRDTIFVSLSTQ 542
 DB 143 IGRKSLMTAAIPNIGLWLAISPAKDALFLYNGRLLGEGVGIIIVTPVYIAEI 222
 DB 90 FGKMTLWVADLFCITGWLAIISLAKDIITWLDNGRFLVGLVGLISVYVVFVYIAETPKHV 149
 QY 223 RGNLGSVNLVSYTFGIFLAYLLGMFIPWRLAVIGALPCTMLIPGLFFIPSPRWLAKVN 282
 Query Match 32.3%; Score 898.5; DB 2; Length 462;
 Best Local Similarity 44.4%; Pred. No. 1.4e-51;
 Matches 189; Conservative 75; Mismatches 153; Indels 9; Gaps 3;
 QY 103 TLIVAGPTQFGTSGFSPTQDMVRDLNLSISFSAFGSLSNVGMVGAISQMAEY 162
 DB 30 TTIIVASFTFGAAGTGYADTWTSSIMSDLSLAQSLFGSLSTFGMIGAFSAKAASA 89
 QY 163 IGRKSLMTAAIPNIGLWLAISPAKDALFLYNGRLLGEGVGIIIVTPVYIAEI 222
 DB 90 FGKMTLWVADLFCITGWLAIISLAKDIITWLDNGRFLVGLVGLISVYVVFVYIAETPKHV 149
 QY 223 RGNLGSVNLVSYTFGIFLAYLLGMFIPWRLAVIGALPCTMLIPGLFFIPSPRWLAKVN 282

```

Db 150 RGAFTHNQQLQNCQGVAVVYFGNLSWRTLAIIGSIPCMQVIGLFFIPSPRLAKG 209
QY 283 LTEDCETSLQVLRGFETDITTEVNDIKRAVASSKRTTISQELNOKKYRTPILLGLL 342
Db 210 RDKCEEVQLKRGKXDIVPEACEIKISVSBAKXKNINIRSLFEKRYAHQLTIGILM 269
QY 343 VLQNLGSLGVLFYASSIFKKAAGVTNSDLATCSLGAIOVLATGVTTMLDRAGRRLII 402
Db 270 LLQQLCGTAGISSYSTFLKAGPP-ARIGMMVLSLIVVPSKLMGLIIVDRWGRPLMT 328
QY 403 STSGMTICLLAVSVVFLKD--NISQDSNSYIITMISLVGIVSFVITPFGMGAIPWLM 460
Db 329 SALGLCSCITLAVAFGVKDPVGGK-----ITPFCFGLSFTMFAFGMALEPWI 382
QY 461 MSEILPVSIKSLGSIATLANWLSFAITMTNMLTWSVGTFLSYMYVSFAITVFWL 520
Db 393 MSEIFPMIDKVLASLVTIANFWFTGWIAFYAFNFWLWSPSGTIIISAIICGATIVTWC 442
QY 521 WVPETK 526
Db 443 LVPETR 448

RESULT 13
Q9MAA4 PRELIMINARY; PRT; 463 AA.
AC Q9MAA4
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Putative sugar transporter.
GN Name=12HI.11;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Town C.D., Benito M.-I., Creasy T.H., Haas B., Wu D.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBSJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL; AC009177; AA270321.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transp.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; Sugar tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR TIGRfams; TIGR00879; SP: 1.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 463 AA; 50148 MW; 82B93947AE79DAC0 CRC64;

Query Match 32.0%; Score 890; DB 2; Length 463;
Best Local Similarity 41.2%; Pred No. 5.3e-51;
Matches 18; Conservative 94; Mismatches 149; Indels 18; Gaps 5;

QY 91 VLRESHVSAPFLCTLIVAGLPIQFGFTSGFSPTQDAMVRDNLISBFSAFGSLSNVGM 150
Db 20 VSEASMMVYLSIIIAVCGSVFQTCVGSAPTQFGIMEELNLSYQFSVFGSILNMGAV 79
QY 151 VGIATSCQAEYICRKCSLMIAIPNIIHGLIASFAKDAFLYKRLLEGFGVIGIISYTV 210

```

```

Db 80 LGAITSGKISDFIGRKGAMRLSSVISAIGWLIILYAKGVPLDFGRFLTYGGCTLSFVV 139
QY 211 PVIYAIISQNRMGALGSVNQISVTPGIGIPLAYLLGMFIWRLLAVIICALPCTMLIGLFF 270
Db 140 PVIYAIISPRKRGALATNLQNFIVIGLASMFELIGAVANWRTALTGVAPCVVLPFGTW 199
QY 271 IPSPRWLAAMNLTECETSLQVLRGFETDITTEVNDIKRAVASSKRTTISQELNOKK 330
Db 200 IPSPRWLEWGVGRHSDFEALQKLRGPQANITREAGEIOEYLASLAHLKATLMDLIDKK 259
QY 331 YRTPLLGLGLLVQLNLSINGVLFYASSIFKKAAGVTNSDLATCSLGAI-----QVLATG 385
Db 260 NIRFVIVGVGLMFPFOQFVGINGVIFVYAOQIFVSAG-----ASPTLGSILYSIEQVLT 313
QY 386 V-TTLLDRAGRRLIIISTSGMTICLLAVSVVFLK-DNISQDSNSYIITMISLVGIV 443
Db 314 LGATLIDRLGRPLMASAVGMLICLLIGNSLFKAHGLAD-----IIPALAVSGVL 368
QY 444 SFVITPFGMGAIPWLMSEILPVSIKSLGSIATLANWLSFAITMTNMLTWSVGT 503
Db 369 VVIGSFSGMGAIPWIMSEIPNLKTAGGLVTVNWLSSWLSVFTFNLWINSFHT 428
QY 504 FLSYMTVSAFTIVFVVLWVPETKG 527
Db 429 FVYGVGVCVLAIFIKLVFETKG 452

RESULT 14
Q94KE0 PRELIMINARY; PRT; 470 AA.
AC Q94KE0
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 05-JUL-2004 (TremBLrel. 27, Last annotation update)
DE At1g08920/F7G19.20.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,
RA Powner L., Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M.,
RA Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M.,
RA Seki M., Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K.,
RA Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Banh J., Bowser L.,
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBSJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL; AF367260; AAK56249.1; -.
DR EMBL; AY133547; AA91377.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR006162; Ppantne_S.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transp.
DR InterPro; IPR005829; Sug_transporter.

```

Search completed: October 13, 2004, 11:36:09
Job time : 66.1256 secs

DR	Pfam: PF00083; Sugar tr; 1.
DR	PRINTS: PR00171; SUGRTNSPORT.
DR	TIGRFAMs: TIGR00879; SP; 1.
DR	PROSITE: PS08650; MFS; 1.
DR	PROSITE: PS00112; PHOSPHOTANTHREINE; UNKNOWN_1.
DR	PROSITE: PS00216; SUGAR_TRANSPORT_1; 2.
KW	Sugar transport; Transmembrane; Transport.
SEQ	SEQUENCE 470 AA; 51299 MW; B251FCBA8419A772 CRC64;
Query Match	31.58; Score 874.5; DB 2; Length 470;
Best Local Similarity	39.4%; Pred. No. 5.8e-50;
Matches 174; Conservative	95; Mismatches 158; Indels 15; Gaps 4;
QY	92 LRESHVSA--FLCTLIIVALGPIQFGTSGFSPQTQAVRDNLNISISFSAFGLSNVGG 149 :
DB	23 INECRIITAVLVFTSVGVCSGFCFCAAGYSVAQTGIINDLGLSLVAQYSMFGSIMTFGG 82 :
QY	150 MVGAIASQWAYEIGRGKSLMIAPNLTIGMLAISPAKDASFLYNGRLLEGFGVGLIISYT 209 :
DB	83 MIGAFSGKVADLMGRGTVMWFAQICFIQGWAVAALAKDSMWLDIGRLSTGFVGLLSYV 142 :
QY	210 VPVVIAETISPNMGEGALGSYNQLSVTFGIFLAVILGMFIEMRWLLAVIGALPTMLIFGLF 269 :
DB	143 IPVIAETIPKHVGAFVFANQLMQSCGLSLFYVIGNFVHWENLALIGLICPALQVVTLF 202 :
QY	270 FIPSPRWLAQNUTDCETSLOVLKGPETDITEVDNIKRAVASSSKRTTISIQELNQK 329 :
DB	203 FIPSPRLILKGWGHEKRASLOSRLGGDDADISEEANTIKEITWLFDGPKSRVMDUFQR 262 :
QY	330 KYRPLLAGICLLVLQNLGINGVLFYASSIKAAQVTNSDLATCSLGAIOLATGVTTW 389 :
DB	263 RIAPSIVIGVGLMLLQOLSSESSGLMYVGVDFKGGFPSS-IGSMILAVIMPRALLGLI 321 :
QY	390 LDLRAGRRIILIISTGMTCLLIAVSVPFLKNISQDSNSYYIL-----TWISLGVISVF 445 :
DB	322 IVERMGRRLPALLASTGCMCFPSLILSFCFR-----SYGMLDELTPITCIGVWGF 373 :
QY	446 VTTFSGMGATPWLMNSEILLPVSTKSLGGSIATLANWLTSAITTNMLMTWSVGGTFL 505 :
DB	374 ISSFAVGGGHPWIINSEIFPMVKVKSAGILVTLAWSGWIVAFAYFMLEWNASCTFL 433 :
QY	506 SYMVVSAPTIVFVVLWPETKG 527
DB	434 IFFTTCGAGIVFIYAVWPETKG 455
RESULT 15	
Q94CI6	PRELIMINARY; PRT; 478 AA.
ID	Q94CI6 AC Q94CI6;
DT	01-DEC-2001 (TrEMBLrel. 19, Created)
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	Sugar-porter family protein 2.
GN	Name=SPB2;
OS	Eukaryotes thaliana (Mouse-ear cress).
OC	Rubidyota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OX	eurosid II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;	
CC	[1]
SEQUENCE FROM N.A.	
MEDLINE=21377429; PubMed=11485201;	
Quirino B.F., Reiter W.-D., Anasino R.M.;	
"One of two tandem Arabidopsis genes homologous to monosaccharide	
transporters is senescence-associated.";	
Plant Mol. Biol. 46:447-457(2001).	
- - SUBCELLULAR LOCATION: Integral membrane protein (By similarity).	
EMBL; AY026255; AAK11721.1; -.	
GO; GO:0016021; C:integral to membrane; IEA.	
GO; GO:0005351; F:sugar porter activity; IEA.	
GO; GO:0005215; F:transporter activity; IEA.	